JOURNAL OF CONSULTING PSYCHOLOGY

VOL. 13, No. 5
OCTOBER
1949

EDITED BY
LAURANCE F. SHAFFER

ASSOCIATE EDITORS: EDGAR A. DOLL · WILLIAM A. HUNT · E. LOWELL KELLY · MORRIS KRUGMAN
BERTHA M. LUCKEY · FRED McKINNEY · CATHARINE C. MILES · CARL R. ROGERS · R. N. SANFORD

CONTENTS

MENTAL DEFICIENCY—SYMPTOM OR DISEASE? VIRGINIA M. AXLINE 313

PSYCHOLOGICAL TESTS IN CLINICAL RESEARCH ROY SCHAFER 328

FACTOR ANALYSIS OF DISCRETE RESPONSES TO THE RORSCHACH J. R. WITTENBORN 335

THE RORSCHACH AS A MEASURE OF INTELLIGENCE W. D. ALTUS AND GRACE M. THOMPSON 341

VALIDATION OF THE ROTTER INCOMPLETE SENTENCES J. B. ROTTER, JANET F. RAFFERTY AND EVA SCHACHTITZ 348

VALIDITY OF SHORT FORMS OF THE WECHSLER-BELLEVUE FOR MENTAL DEFECTIVES M. COTZIN AND J. J. GALLAGHER 357

ANALYZING AND EXTENDING THE USE OF BELLEVUE-WECHSLER VOCABULARY RESPONSES REVA A. GERSTEIN 366

GROUP PSYCHIATRIC SCREENING OF RECRUITS N. G. HANAWALT AND R. A. KIDD 371

INTELLIGENCE AND SCHOOLING WHEN LITERACY IS CONSTANT WILLIAM D. ALTUS 375

WORRY PATTERNS OF AMERICAN PSYCHOLOGISTS W. A. KERR, H. L. NEWMAN AND A. R. SADEWIC 377

NEW BOOKS AND TESTS 381

AMERICAN PSYCHOLOGICAL ASSOCIATION INCORPORATED

JOURNAL OF CONSULTING PSYCHOLOGY

ESTABLISHED 1937

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

THE JOURNAL OF CONSULTING PSYCHOLOGY is devoted to the fields of clinical and consulting psychology, counseling and guidance. It publishes the results of research bearing on these areas of psychology; reports, discussions, and case studies in clinical psychology; theoretical contributions pertinent to psychological diagnosis and counseling; and articles on the professional practice of psychology for human welfare.

THE JOURNAL OF CONSULTING PSYCHOLOGY was founded in 1937 by the Association of Consulting Psychologists, and from 1938 to 1945 was the official organ of the American Association for Applied Psychology. With the amalgamation of the latter society with the American Psychological Association in 1945, the JOURNAL was continued as the official publication in its specialty.

EDITORIAL OFFICE

Manuscripts submitted for publication, books and tests for reviews, and all correspondence on editorial matters should be sent to:

DR. LAURANCE F. SHAFFER, *Editor*,
525 W. 120th St.,
New York 27, N. Y.

Contributions relevant to the field of specialization of the JOURNAL are invited. Manuscripts should be typed with wide margins, and double-spaced throughout, including titles, quotations, footnotes and references. The style of all articles, figures, tables, and references must follow the form given by John E. Anderson and Willard L. Valentine. The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345-376. Copy for figures and illustrations must be prepared for reproduction.

Proof will be submitted to the author shortly before publication, and an opportunity to order reprints will be given at that time. Changes made by the author on the proof are chargeable to him.

BUSINESS OFFICE

Communications concerning subscriptions, change of address, claims for nonreceipt of a number and other business matters should be sent to:

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1515 Massachusetts Ave., N. W.,
Washington 5, D. C.

THE JOURNAL OF CONSULTING PSYCHOLOGY is published bi-monthly, in February, April, June, August, October, and December, the yearly volume comprising approximately 440 pages. The subscription per year is $5.00, foreign $5.50; single numbers $1.00.

Copyright 1949 by the American Psychological Association, Inc.

Printed at 23 West Colorado Ave., Colorado Springs, Colo., and entered as second-class matter at the Post Office at Colorado Springs, Colo. Acceptance for mailing at special rate of postage provided for in section 538, Act of February 25, 1925, authorized May 27, 1940.

Volume 13 OCTOBER, 1949 No. 5

JOURNAL OF
CONSULTING PSYCHOLOGY

MENTAL DEFICIENCY — SYMPTOM OR DISEASE?

By VIRGINIA M. AXLINE
TEACHERS COLLEGE, COLUMBIA UNIVERSITY

FOR a good many years the term "mental deficiency" has been used very loosely to indicate a fixed state of mental ability — or a definite limitation of the individual's capacity. The degree of deficiency was quite often determined by standardized tests. Once a child was properly diagnosed and had a label duly attached to his record he was more or less doomed to the kind of treatment that was based on a limited prognosis of his mental and social development. His "educational" opportunities were restricted to those experiences that were considered best for the *training* of that limited capacity. There were few efforts made to attempt any treatment that might seek to determine the nature and flexibility of the capacity of such individuals and to provide educational experiences that might increase the capacity. There were few investigations into the highly individual nature of the "mentally defective" child's behavior. There was a marked tendency to do quantitative studies rather than finely detailed qualitative studies that might throw a more penetrating light on the dynamic nature of mental "deficiency."

Recently, however, there have been an increasing number of psychologists and educators who have raised questions in regard to the old definition of mental deficiency. Is it a disease? Or is it a symptom? What is the nature of mental deficiency? What are the dynamics of such behavior? How do they compare or deviate from so-called normal behavior? Does an IQ of 68 doom the individual to ever-lasting feeblemindedness? Does an IQ of 150 guarantee a life of psychological achievement? These are questions that have no cut and dried answers at this moment. Our psychological understanding of the nature of mental deficiency is in the pioneer stage. The holistic concept of individual psychology includes the child's physical, emotional, social, and intellectual behavior as a dynamic, interacting process.

Objective, detailed studies of individual child behavior in the kind of situation that provides natural, spontaneous self-expression of the child's inner life present some observational material for thoughtful consideration.

The following study is presented as exploratory material to raise again the question: What is mental deficiency? This material is *not* presented as an attack upon the validity of IQ tests—or as propaganda for any specific type of

treatment. It is not claimed that the treatment "raised the IQ". In those instances where there were marked changes in IQ scores after the completion of therapy the explanation that seems apropos is that the child was freed from emotional constraint and could thus express his capacities more adequately.

Data for this study consisted of the records of fifteen six- and seven-year-old children who had been referred for play therapy because they were "behavior problems," "emotionally disturbed children," "speech problems" — according to the parents' preliminary interviews regarding these children. Each child was seen individually by the same therapist for a series of play therapy interviews—ranging from 8 contacts to 20 contacts. This study was drawn from the accounts of the therapy sessions as recorded by trained observers—these reports being as nearly verbatim and objectively descriptive as stenographic reports can be. The therapy experiences were not planned to furnish experimental evidence for a study of "mental deficiency" as observed thru play therapy experiences. Rather, the case reports were drawn from the files—selected on the basis of IQ ratings and the age of the children—grouped and analyzed at a later date because the material seemed to lend itself to this study of mental deficiency. It is presented as a pilot study only—suggestive of ways of doing more rigorous and detailed study of the dynamics of child behavior that becomes objectively available from recorded play therapy interviews.

There were three groupings — each group containing five children. The first group was composed of children who

TABLE I
SUMMARY OF DATA ON THE FIFTEEN CHILDREN IN GROUPS I, II AND III

Group I						
Child	Sex	Age	Pre IQ	Post IQ	Number of Interviews	Remarks
Lenny	Male	6.4	66	67	8	Speech Problem
Orville	Male	6.2	73	70	15	50% Vision
Charlie	Male	7.4	69	69	15	Speech Problem
Lucy	Female	6.1	70	74	20	Spastic
Charlene	Female	7.3	72	76	15	Speech Problem
Group II						
Child	Sex	Age	Pre IQ	Post IQ	Number of Interviews	Remarks
Jim	Male	6.7	65	85	10	Deaf
Paul	Male	6.5	74	89	12	30% vision
Junior	Male	7.6	72	96	17	"Non-talker"
Babbette	Female	7.0	71	90	18	Spastic
Sissy	Female	6.9	74	100	12	60% vision
Group III						
Child	Sex	Age	Pre IQ	Post IQ	Number of Interviews	Remarks
Joseph	Male	7.6	110	112	17	Post-polio
Henry	Male	7.0	114	112	14	"Non-talker"
Harold	Male	6.8	106	106	10	Speech Problem
Nancy	Female	6.5	105	100	12	"Non-talker"
Eleanor	Female	7.3	106	111	12	50% vision

showed no appreciable change in IQ scores after therapy—both pre and post-tests indicating "low" intelligence. The second group was composed of children who showed noticeable gain in IQ scores—the pretest indicating "subnormal" intelligence, the post test indicating "normal" intelligence. The third group was composed of children who showed "average" intelligence according to pre- and post-tests. The test given to all the children was Stanford Binet (L and M). Table I gives the individual details of Group I, Group II, and Group III.

Brief summaries from the preliminary interviews as given by the mothers or mother substitutes describe the children's behavior as the parent interpreted it; psychologist's notes describe behavior during test sessions; therapist's reports describe behavior during play therapy.

GROUP I-CHILDREN WHOSE IQ STAYED LOW

"*Lenny* has always been queer. He doesn't talk. He just jabbers and makes odd noises. He imagines things I think. He can say a few words. He thinks he sees snakes. And he acts scared. Trembles. Cowers in a corner. Doesn't seem to understand anything we say to him. Doesn't play at all. Stands around and sucks his fingers. Then suddenly screams. He has a funny lopsided walk—sort of half-trot half-walk way of getting around. He whimpers—but never really cries. He can't do anything for himself." This was the mother's report.

Lenny cooperated with the psychologist—seemed to listen to the directions for the test—secured a rating of 66 on the first test—67 on the second test. The psychologist noted a more extensive vocabulary than his mother had reported, altho it was meager. He seemed quite fearful during the test, sucked his fingers a lot, screamed occasionally, whimpered, but did seem to comprehend most of what was said to him.

The play therapist noted the evidences of insecurity, anxiety, and fear during the first three interviews. He stood in the corner of the play room whimpering. There were many toys in the room including a sand box, little cars and airplanes, families of dolls, baby doll, nursing bottles, clay, paints, crayons, doll house, sink and running water, toy dishes, guns, drum. Lenny did not play with any of the toys. He did not seem to look at them. He chewed his fingers, looked at the wall, screamed and stiffened, kicked his foot weakly against the wall. During this time the play therapist stayed in the room with him, said during the first interview that he could play with the toys if he wanted to. Then she let Lenny decide how he wanted to use that time. She smiled at him, sometimes spoke to him, to indicate he could play or not—as he wished. For three interviews Lenny behaved in this manner.

During the fourth interview he went to the table and tore all the drawing paper to shreds screaming all the time. During the fifth interview he threw the sand wildly around the room—screaming all the time in a thin feeble manner. During the sixth interview he sat down at the table and broke all the crayons and dropped them on the floor, crushed them under his foot. During the last five minutes of this hour interview he scribbled aimlessly with a bit of broken crayon on the table top. During the seventh interview he sucked on the nursing bottle. During the eighth interview he tore apart the big rag doll and threw sand at it. He screamed, kicked it, threw water on it. This was his last interview with the play therapist. He was given the second Binet test immediately following this interview by the psychometrician. Her report noted that he was more relaxed than in the first testing session—that he even laughed a little.

Lenny was then placed in an institution for the feeble-minded—and therapy was discontinued.

The one factor that stands out above all others seems to be the sudden initiation of play on his part. The therapist had made no attempt to draw him into any activity—feeling that it was *psychologically more important for Lenny to initiate the activity that to be playing* because of her gentle persuasion. The fact that Lenny spent such long periods of time on such activity is interesting. The aggressive violence of the play

and especially what seemed to be a continuity of activity in the last bit of doll play suggests that Lenny might have been beginning to express some of his feelings thru the medium of these toys. If his mother's initial report is accepted as accurate Lenny could be said to have made amazing progress in a few short interviews. One can only speculate on the replacement of the whimpering fearful anxiety that was first observed by the violent aggression that was unleashed. It suggests that the inner life of this child was in marked contrast to the feeble and futile passivity that had generally been noted.

Orville's mother described him as "wild." She said he was "either whimpering like a very small baby or having violent temper tantrums." She said he seemed erratic, had "no attention span at all," could not get along with other children; ate paper, crayons, matches, dirt; ran and hid when any visitors came to the house and would not come out from under the bed or the closet—wherever he happened to hide. She said he could not converse with her or any one else—that he did talk—but what he said was never in response to what anyone said to him—but rather repetition of what he heard. She said she was planning on sending him away to school for "slow learners."

The psychologist reported difficulty in getting and holding his attention. He tried to keep the test materials—snatched at things—banged his head down on the table—struck at her, then whimpered. In spite of all this however she was able to complete the test.

During the play therapy sessions one thru five he dug in the sand box and systematically buried all the toys. Then he got out the hammering set and banged on it. Each time he came into the playroom he followed the same pattern of behavior. During the sixth interview he became very aggressive—tipped over the chairs and the doll house, threw water on the floor, threw sand on the floor. During the seventh interview he scribbled with crayons and later went back to the sand box and buried all the toys. During the eighth interview he crawled around the floor like a baby, drank from the nursing bottle—then suddenly became aggressive and upset the furniture, doll house, and threw water on the dolls. During the ninth and tenth interviews he messed with finger paints—smeared it all over his arms, face, the table, the chair. This was the first interview that he seemed to enjoy. He laughed—was visibly relaxed and did not want to go when the time was up. He made a few remarks directly to the therapist—such as "Me like to do this." "This fun." "Me put it on me." During the eleventh interview he played in the water at the sink—splashing the water—letting it run in and out of the sink. Then he ran over to the sand box and dug a big hole in the sand, came over and got a cup of water, poured it in the hole and "buried" the water. During the twelfth interview he took all the furniture out of the doll house, then played in the water, ended the session by "burying" some water in the sand box again. During the thirteenth interview he scribbled with the crayons, finger painted, then played in the water. He again "buried" some water in the sand. He would say to the therapist "Dig a hole and bury it big." He buried the crayons during the last few minutes of this session. He was told he would come only two more times.

During the fourteenth interview he played with the blocks, the crayons, the doll, the water. He talked to the therapist quite a lot—was more relaxed—seemed happy. At the conclusion he was told again that he would come only one more time. At this statement he began to vigorously bury all the toys he had used in the sand box.

During the fifteenth and last interview he spent the entire period burying everything in the room that he could possibly put in the sand box—even trying to bury the doll house.

At the conclusion of this interview he was sent away to the private school for "slow learners."

Charlie was described as a "severe behavior problem." His mother said he was "disobedient, sassy, defiant or else he was sulking and hiding away by himself."

The psychologist reported difficulty in establishing rapport, but once she did get his attention he tried very hard to do what she told

him to do and became very affectionate toward her—trying to kiss her hand and hanging on to her.

During Charlie's fifteen play therapy sessions he seemed to be doing exactly the same thing week after week. He would come into the room and upset everything—then he would select the little shovel and go around the room sprinkling sand on everything. The interesting factor in Charlie's case was the conversation that he carried on with the therapist. It followed a pattern too. He would say "I'll kill this. I'll kill this. I'll knock down the place." He took great pleasure in kicking the rag doll and in squeezing its neck. During the last fifteen minutes of the fifteenth interview Charlie "fixed up the room." He straightened up the furniture, swept the floor, smoothed the sand. Then he sat down on a chair and looked silently at the therapist with tears running down his face. He seemed to be unable to communicate verbally the feelings that overwhelmed him. The play therapy interviews were terminated because Charlie was being sent to another town to stay with his father and paternal grandmother. Charlie's parents were divorced—and the custody of the child was "shared" by the parents. The mother had explained at the beginning of the therapy experience that he would leave at this date because she "had had Charlie more than her share anyhow."

Lucy was described as "a total blank" by her mother. Lucy's talk was so garbled it was almost unintelligible. The mother said that there were moments when Lucy seemed almost normal—but that more often she seemed "to completely withdraw and live in a world of fantasy." She had no friends—chewed her hand, had "terrible nightmares," wet the bed, cried, seemed very tense and stiff in all her movements.

The psychologist reported that Lucy had been very tense and anxious during the tests. She stared at the psychologist and reacted at an extremely slow rate. She showed only minor symptoms of spasticity.

During Lucy's first four play interviews she sat on a chair looking either down at the floor or jerking around and staring blankly at the therapist. The therapist did not attempt to push Lucy into any activity—but tried to help Lucy relax and feel at ease by keeping external pressures away from her during these play periods. During the fifth interview Lucy picked up the nursing bottle and sucked on it during the hour. She looked around the room—smiled briefly at the therapist. During the sixth interview she climbed awkwardly into the sand box and shoveled the sand somewhat aimlessly. During the seventh interview she began to toss the sand around aimlessly. At the beginning of the eighth interview she came into the room seemingly with a purpose. She walked up to the sand box, climbed in, and began to throw handfuls of sand around the room—yelling feebly while she did—glancing at the therapist—and continuing when she saw no reaction from the therapist.

During the ninth, tenth, eleventh and twelfth interviews she entered the play room with gusto—climbed into the sand box and began to throw the sand. Her yells and screams became louder and freer. She laughed—kicked the sand—jabbered constantly. During the thirteenth interview she upset the chairs, brushed all the doll house furniture out on the floor, threw the doll furniture against the wall—climbed into the sand box and ran her hands through the sand— pressed her face down against it, smiled at the therapist, crawled over and got the nursing bottle, dragged a chair up close to the therapist, snuggled up, patted the therapist's hand, rubbed her cheek up against the therapist's arm, sucked on the nursing bottle. During the eighteenth interview she sat down at the table and scribbled with the crayons on a piece of paper. She gave it to the therapist, smiled, patted her hand, jabbered—apparently trying to communicate something to the therapist. Then she went to the doll house and "arranged" the furniture. She got the rag doll and sat down and hugged it, pretended to feed it with the nursing bottle—came over close to the therapist and clutching the doll in one arm, holding the nursing bottle between her knees, took the therapist's arm and placed it around her (Lucy's) shoulders—then picked up the nursing bottle alternately fed the doll, drank from it herself, and smiled up at the therapist. She said quite distinctly "Baby." "Baby." "Me baby." During the nineteenth interview she played with the doll again—this time placed the doll's high chair up to the table and carefully fed the doll. Then she offered the therapist a cup of sand, poured one for herself. She rocked the doll—placed the doll in the sand box and put the shovel in the doll's hand. She played in this manner during the entire hour—drawing the therapist into the

play at intervals—once placing the doll in the therapist's lap, removing it and climbing up on the therapist's lap herself, getting down getting the nursing bottle, drinking from it. During the twentieth and last interview Lucy had been told it would be her last time. The family was going to the country for the summer and Lucy's play sessions were to be dropped. Lucy sat close to the therapist hugging the doll, then suddenly she jumped up threw the doll across the room, ran around the room screaming and upsetting the chairs and table and doll house. She threw the sand in wild handfuls. When some of it hit the therapist she ran over and brushed it off and grabbed the therapist's hand—leaned down and set her teeth on the edge of the therapist's hand but kissed it instead, ran into the corner and buried her face in her arms, came out and began to straighten up the room, picked up the nursing bottle and sat down beside the therapist until the time was up.

Charlene's mother decribed her as "a disgrace to the family." She said her other children were "quite bright" and they were all ashamed of Charlene. Charlene never talked to anyone, altho the mother said that Charlene *could* talk but seemed to enjoy "aggravating the family" by her refusal to answer them. She said Charlene did not play—"just stood around — sometimes deliberately scratched and ruined the furniture—and if she had a chance she would mark on the wall paper, cut the table cloth or the curtains—or anything that belonged to others in the family, but never her own things."

The psychologist reported difficulty in establishing rapport—finally Charlene settled down, answered the psychologist's questions, followed the directions as best she could.

During the first play therapy interview Charlene spent the time looking at all the toys and materials in the play room. She flitted from one thing to the next—picked up the toys, examined them, dropped them on the floor, "accidentally" knocked over a chair. During the second interview and the nine following interviews Charlene spent her play sessions upsetting the doll house, throwing sand on the floor, knocking the dolls around, pounding them with a hammer. She muttered and whispered to herself most of the time. Sometimes she talked to the dolls. "You do what I tell you. I'll beat you. I'll send you 'way." Occasionally she would address a remark directily to the therapist—such as "I got a doll at home but I don't like it." or "I'll throw sand on your old floor." or "I'll break this old doll furniture."

During the eleventh interview Charlene finger painted the entire hour—became completely absorbed in it—used only red finger paint, lots of it—and did not say a word during the hour. When it was time to go she protested —but finally accepted the time limit, said goodbye and left.

During the twelfth, thirteenth, and fourteenth interviews Charlene finger painted. She used both hands and her arms. She traced patterns in the finger paint. At first they were very chaotic and disorganized. Later they had more form and balance. During the fourteenth session she was very relaxed, talked to the therapist occasionally—said she liked the paint —said she didn't want this next time to be her last time.

During the fifteenth interview she played with the doll house, drank from the nursing bottle, hammered on the peg board, undressed the doll and put it to bed, then climbed into the sand box and played with the little cars in the sand. At the conclusion of the fifteenth interview Charlene was sent away from home to attend a special school for slow learners.

In each of the above cases presented here in thumb-nail fashion the reader has a glimpse of the way these children expressed themselves in the play situation. In each case the mother was interviewed twice by the play therapist—once at the beginning of the therapy—again at the conclusion. In every case the mother expressed shame, disapproval, and rejection of these children, emphasized by the fact that they were sending the children away. Certainly it would not be stretching the point to say that each of these children were victims of extreme emotional deprivation. Their primary difficulty seemed to be in their ability to express their feelings and emotions in their daily lives without increasing the threat that their families already presented to them. They were

not able to communicate clearly to others the things that were uppermost in their lives. They could only be victims of their own emotional storms. They were not able to establish a satisfactory relationship with others. And it is a fallacy to assume—if anyone does—that these children do not suffer acutely from this emotional deprivation—or to assume that they are unable to establish relationships with others because their ability to communicate with others is one of their handicaps.

In each of these cases the therapist felt that the therapy was incomplete—that what had happened was sufficiently satisfactory from the standpoint of a favorable prognosis to have followed thru on each of these cases. However, it was impossible to do so.

Some of the activities in which these children engaged hint at a symbolic expression thru the medium of these play materials that is very like the use other children with presumably more intellectual capacity make of it. The continuity of each child's activity during a series of play therapy contacts poses a question of the psychological significance of such a therapeutic experience for children.

These five children are considered acceptible examples of mental deficiency. Is it a disease? Is it a symptom?

CHILDREN WHOSE IQ IMPROVED

Thumb-nail sketches of the children's experiences in Group II are presented for comparison with Group I.

The mother's preliminary report concerning *Jim* was so filled with emotionalized attitudes of rejection that it was difficult to extract the descriptive terms she used to describe Jim. They included such expressions as "unbearably stupid," "uncontrollably wild," "heartbreakingly bad," and "like a little animal."

The psychometrician reported Jim as "an extremely fearful child—looking nervously about him. Tense. Nervously biting his nails," "Made peculiar gutteral noises throughout the test," "grimacing," "whimpering," "cowering in his chair" when she asked him a question—but watching her lips as though trying to understand. The score on the first test was 65.

The play therapist saw Jim for ten times—knowing ahead of time that at the conclusion of the tenth interview Jim would not return. He was being sent away to a "special school" several hundred miles from home.

When Jim first came into the playroom he stood just inside the door whimpering and chewing his fingers and playing with the door knob. The play therapist motioned to the toys but Jim hid his face in his hands and stayed by the door during the hour. During the second session, Jim stayed by the door and behaved in the same fashion for ten minutes, then he went swiftly around the room examining all the toys closely, dropping them on the floor when he finished looking at them. During the third interview he came into the room swiftly, repeated the behavior of examining all the toys and dropping them on the floor, then he began to pile all the furniture and toys in a big pile in the middle of the floor. During the fourth interview he began to pile the toys and furniture in the middle of the floor and then got in the sand box and threw shovels of sand at the pile of furniture and toys. During the fifth interview he came in and went to work clearing the room of all the toys—hiding them in the cupboards—under boxes—in the sand box. He worked swiftly—making peculiar grunting noises all the time. During the sixth interview he repeated the same behavior but during the last ten minutes of the period selected the rag doll and took all its clothes off and buried the clothes—then he hammered the doll viciously, jerked its head off, bit it.

During the seventh interview he painted at the easel—streaks of black paint up and down on the paper. During the eighth interview he came into the room and piled all the furniture in the middle of the floor, making a barricade of it and hiding behind it, then he played with the doll house—first piling all the furniture in a pile in the middle of the rooms in the doll house and covering it with sand and then knocking all the furniture out on the floor—and straightening up the doll house and putting the doll family in bed. Then he came over to the therapist, sat down very close to her and took her hand, smiled at her. He was very re-

laxed—leaned back in the chair and looked at the therapist. During the ninth interview he played with the clay for ten minutes, held the doll in his arms, kissed it, fed it from the nursing bottle, put it to bed and covered it very gently. Then he got out the toy dishes and set the table for two and said to the therapist—"Come"— motioning across the table from himself. The therapist took the chair across the table from him. He smiled and sat there looking at the therapist until it was time to go.

During the tenth interview he fixed the place across the table from himself for the therapist and sat down across from her. He got a paper and some crayons and printed his name—JIM. He gave it to the therapist, smiled. Then he took another piece of paper and wrote 1 2 3 5 7 8 9 and handed that to the therapist and smiled. Then he took another piece of paper and drew a picture of a house and a child outside. The "child" was the typical drawing of face, arms, legs that young children so often draw. Then he drew a big face with the mouth going down and tears on the face and handed it to the therapist. Then he picked up the doll and kissed it. When it was time to leave he said "Bye-bye" — waved, smiled, and left.

Paul was described by his mother as "a stubborn, aggravating, slow child—always saying either 'Huh?' or 'Why?' " He wore glasses—and during fits of temper he would frequently throw his glasses on the floor and break them by stamping on them. He had "nightmares and temper tantrums." He "wet the bed just for meanness." He sometimes "refused to eat for two or three days at a stretch." He "deliberately broke things—like his dishes during meal times—would deliberately upset his milk." He spit at people—and was "surly and defiant." The mother summed up by saying wearily "No matter how hard I try I can't think of one nice thing about him."

The psychometrician had difficulty establishing rapport. Finally she did gain his attention long enough to give the test and got only a moderate degree of cooperation from Paul.

During the first interview Paul quickly sized up the play-room situation—and went around the room hitting all the toys with the peg-pounding hammer. During the second interview Paul spent the entire period tearing the rag doll apart and picking out all the cotton. During the third, fourth, fifth, and sixth interviews he spent his time throwing the sand out on the floor, at the wall, at the toys—and at the therapist, but that was limited.

He accepted the limitation to "throw it at a doll—or any place else in the room" but *not* at the therapist. During the seventh interview he rolled the clay in little balls and threw them at the doll house. Then he took the family of dolls and laid them in a row on the table and covered them with mounds of clay—yelling at the dolls "How you like that? How you feel now?" Then he got the hammer and pounded them as hard as he could. During the eighth interview he played with the gun—"shot" every toy, "shot" the therapist—yelled —"You *dead*!" "*You* dead!" During the ninth interview Paul came in very quietly, sat down at the table and began to draw. Suddenly he threw down the crayon, buried his face in his arms and cried. Then he got up and went over to the doll house and pushed it over—kicked the spilled-out toy furniture—took all the dolls, dug a big hole in the sand box and buried them. Then he took his chair over in the corner farthest from the center and sat down with his back to the therapist—doing nothing until it was time to go.

During the tenth interview Paul painted—streaks of color going in every direction. During the eleventh interview he played with the little cars in the sand box—and with the toy soldiers. He talked to the therapist for the first time—telling her what he was doing. "I'm going to have a war. These soldiers are going to go kill these Germans. These soldiers are going to kill these Japs. These are good soldiers. They only kill the right people. This little old car is going to come out and give them a big ride and take them across the world. This soldier is lost and he's afraid." All during this play he described the imaginative activity. During the twelfth interview he painted a house, told the therapist it was his house—and then tried to make a house out of clay. After working on it for about twenty minutes he gave up the idea and returned to his painting. Therapy was interrupted at this point because the family moved out of town.

Junior was referred for play therapy because he had "regressed back to an infantile state" according to his mother. He had stopped talking, jabbered like a baby, crawled on his hands and knees in-

stead of walking, ignored everyone—behaving as tho he did not see them.

The pretest had been secured by a nursery school psychologist before Junior had regresser to this state. When Junior was led back to the play room— (he would walk under compulsion—if his mother—or some adult took his hand and said *"Walk!"* with firmness), Junior sat down on the floor—and stayed there all during the first interview — wiggling around, sucking his thumb, jabbering — but seeming to enjoy it immensely. Occasionally he would steal a glance at the therapist. The therapist talked to him about what he was doing—"You want to sit on the floor and suck your thumb, hm?" When he shook his head—"And you shake your head like that—" When he peeked at the therapist "You'll take a little look, hm?" Nothing very profound about the therapist's remarks, but intended to convey to Junior acceptance of him as he was at that moment and to attempt to open some channel of communication between them. During the second interview Junior was led back to the playroom and he climbed in the sandbox without help or suggestion. He spent the period playing with the sand—filling little dishes and cups with sand—pouring it out—filling them again. During the next ten interviews he seemed to do the same thing each time he came. At the beginning of the twelfth interview Junior walked back to the playroom alone—with decision and pep to his walk. He dragged out the blocks and started to build some structure with them. He looked directly at the therapist and began to talk to her briefly. "It'll be my house," was the first thing he said. Then he said "What is your name?" The therapist told him. Then he said "*I* am Junior." There followed a long silence while he built his house. When he had completed it—and placed the big block across the top for a roof he said to the therapist "I *like* my house." Then there were no other comments from him for the rest of the period. During the thirteenth, fourteenth, fifteenth, sixteenth, and seventeenth interview Junior played with the blocks, the sand, and the little cars. He had resumed his talking outside the playroom as well as in it. He no longer crawled around the room at home. He was showing an interest in other people— talking to them—behaving in a very satisfactory and happy manner. The therapy was discontinued because there seemed to be no further need for it.

Babbette was referred for play therapy because she was "a severe behavior problem." She was "always breaking things deliberately, tearing her clothes, scratching and hitting at other people," "yelling at the top of her lungs." Her degree of spasticity was not very marked. There was a noticeable lack of coordination that made her appear very slow and clumsy. Her speech was not clear.

The psychometrician reported "not much cooperation"—"difficulty in holding her attention" —"Babbette was inclined to tease and hide the materials during the test—and seemed to try to divert the attention of the psychometrician."

The play therapist reported extremely aggressive behavior in the play room—repeated attempts to attack the therapist, followed by repeated attempts to hug and kiss the therapist. She threatened to break every object in the room—including the windows, but accepted the limitations to break certain things specified for breaking. She threw the doll furniture against the wall—and tore the heads off all the dolls. Her behavior was cyclonic in style. She screamed, yelled, laughed hilariously. This type of behavior continued for thirteen interviews—the only change seeming to be a lessening of the intensity of the feelings she was expressing. She talked most of the time, but her conversation was unintelligible to the play therapist. During the fourteenth interview Babbette was very different. She was more relaxed. She "played house" with the doll. Her speech was clearer. There was noticeable improvement in her coordination. During the fifteenth, sixteenth, and seventeenth interviews she used the finger paints. This was a new medium of expression to her. She was completely absorbed in playing with the oozing paint—and moved her hands around the paper slowly and deliberately — experimented with the paint—laughed and talked to the therapist about it—said it was "fun." During the eighteenth and last interview she played with the doll—and gave a "party"—put the doll to bed. Then she tried to put away all of the toys —fitting them carefully into the cupboards. The therapy had been concluded because of the summer vacation plans of the family. The parents felt, at the end of the summer, that Babbette no longer needed therapy.

Sissy was referred because of "temper tantrums," "bed-wetting," and "an-

ti-social behavior." The mother was very defensive about Sissy and said she "didn't think any one understood Sissy." "However," she added "altho I feel sorry for Sissy and the three older children are ashamed of her I try very hard not to be. She was a great disappointment to me. She's my baby—and my only girl. I always wanted a girl. And Sissy breaks my heart."

The psychometrician reported that Sissy was "shy but cooperative."

The play therapist reported that Sissy did not in any way live up to the description that preceded her into the play-room. During the first interview she looked around the room, carefully studied the toys, settled down at the doll house and played with the family of dolls. She talked for the dolls. First they planned "a picnic"—then "a long trip"—then "a big party." She smiled at the therapist occasionally as she played.

During the second interview she resumed the doll play, but this time she occasionally spoke to the therapist, telling what the dolls were doing. During the third interview the peaceful doll play was interrupted by violent play of destruction. The house and family "burned up"—a storm came and blew the house over. She hurled the toy furniture around the room—explaining that *she* was the storm. During the fourth interview the toy soldiers moved into the house and took possession of it. The doll family was killed by the soldiers. During the fifth interview she played with the soldiers in the house and "bombed" it with handfulls of sand and blocks. During the sixth interview she "killed off all the soldiers." During the seventh interview she painted. During the eighth interview she played with clay. During the ninth interview she played with the baby doll—fed it—changed its clothes, put it to bed. Then she drank from the nursing bottle and told the therapist she liked it. During the tenth interview she had a party for the doll and asked the therapist to come. The therapist sat at the table across from Sissy. During the eleventh interview Sissy played in the sandbox—"cooking for a restaurant." During the twelfth and last interview Sissy played with the doll, the paints, the doll house, the soldiers and the sand. It was as though she was giving a review of her play sessions. Therapy was terminated because Sissy seemed to have achieved a satisfactory adjustment. She no longer had temper tantrums nor wet the bed. She was playing with other children and getting along well with them.

As one studies the summaries of these cases one notices that Jim, Paul, Babbette, and Sissy were rejected children. Junior's mother did not seem to reject him. The pattern of play displayed by these children is not too unlike that of the first group. There was more movement in these cases. The thread of continuity appears in these cases too. The problems of communication with others and the establishment of satisfactory relationships with others is also apparent in this material. The symbolic expression of feeling thru the toys and the play seems to be more marked. In both groups the children initiated play activity when given an adequate opportunity to do so. In both groups there was an outpouring of negative feelings and destructive play followed by more outgoing and positive behavior. The play expression of these children, coupled with the reasons for referral and the changes during therapy points up the question—What is mental deficiency? —Symptom or Disease?

CHILDREN OF AVERAGE INTELLIGENCE, WHICH DID NOT CHANGE

The play therapy experiences for the children in Group III were conducted in a children's home. The pretest score was obtained from the school files. The post-test given in most cases the week after the completion of therapy by the therapist. These children were all victims of broken homes—the parents had separated and placed the children in the Home. Joseph, Henry, and Eleanor were up for adoption. Harold was deserted by his parents. After placing him in the home and making the initial payment for his care they had disappeared. Nancy's mother lived near the home but seldom went to visit her. Quite obvious-

ly these children were rejected children—emotionally deprived to a tragic degree—insecure, threatened by a shifting, unpredictable world. These children were not "mentally defective" according to the tests. They would represent average or above average intelligence according to clinical standards. These children had been given the pretest in school—a short time before their homes had been broken and they had been sent to the orphanage. Reasons for including Group III in a discussion of mental deficiency—symptom or disease?—will be obvious if we consider it as pertinent to the question *Is it a symptom?* For here is some evidence that rules out mental deficiency as the cause of this behavior. Here are five children known to be rejected, insecure, anxious, and fearful—children who live in the shadowy chill of being unwanted and unloved—children who are suffering from what might better be called *emotional deficiency*.

How important is that deficiency in determining the behavior of the child—socially, emotionally, and intellectually?

The referral cards were filled out by the "house mothers" in the orphanage.

Joseph's card read as follows: "This child is the most disobedient, defiant, belligerent, antagonistic trouble-maker in the home. He seems utterly unable to understand correction. He is a menace to the other children—fights everyone."

When Joseph first came into the playroom, he snarled at the therapist, said he wouldn't stay—that he didn't *want to play*, that he would break everything in the playroom. Then Joseph began to carry out his threats. He tore into the activity with gusto. He upset the tables and chairs—threw the toys over in the corner—threw the sand all over the floor, broke the nursing bottles, screamed, swore, spit, urinated on the floor, dumped the paint on the floor. This kind of behavior continued for the entire 45 minute period. When he was told the time was up and asked if he would like to come back again he said *no* and swore at the therapist. Then he said *yes*, he would come back and do the same thing all over again. Joseph returned for a second interview and repeated the behavior of the first interview. The only change was at the end of the interview when he asked the therapist if he could come back again.

Joseph came for his third interview. This time he played with the gun, built a barricade in the middle of the room out of the furniture and hid behind it—shooting out at the therapist, cursing her, calling her "mom" and "pop" and other very colorful epithets. During the fourth interview he hammered on the pegboard, screamed, yelled—grabbed up the toy telephone and called "mom" and screamed filthy names at her. Then he threw the sand around the room. He spied a big rag doll—placed it on the table said he was going "to torture her"—pounded the doll with the hammer—got a pencil and dug out the eyes, poured the red paint on the doll and said "She bleeds to death." During the fifth interview he played with the family of dolls—filled a basin with water and "drowned" all of them screaming and laughing. "Go ahead and drown. See if I care!"

During the sixth interview he got out the toy soldiers and arranged "camps of enemies" in the sandbox. He said *he* was a Jap and a German and *he* would kill every American he saw. During the seventh, eighth, ninth, tenth, eleventh, twelfth, thirteenth, fourteenth, and fifteenth interviews he played with the soldiers devising ingenious battles. There seemed to be a slight change in his play. There seemed to be a change from violent bombings and screaming to more complicated strategy as he played with the soldiers and tanks and cars.

When he came in the room for his sixteenth interview he got out the nursing bottle, drank from it, crawled on the floor, played baby, came over and sat close to the therapist. During the seventeenth interview he brought a scrap book and showed it to the therapist. He was collecting pictures of airplanes. He also brought some airplane pictures which some other boy has "thrown out" and he spent the time pasting these pictures in his scrap book. The house mother had reported marked improvement in his behavior and had requested that his case be discontinued and another child seen since the therapist's time spent at the orphanage was limited. When asked how he felt about coming back just three more times he replied by saying he would rather go swimming with the other boys. His case was terminated at the close of the seventeenth interview.

Henry's card read as follows: "I don't believe Henry is real bright. He stands around and chews his fingers and whimpers. He never talks—never answers—doesn't seem to comprehend anything. Perhaps he is too hopeless to bother with."

Henry came in the playroom and acted according to the card-description. He stopped just inside the door. The therapist explained to him that he could play with the toys if he wanted to—whatever he wanted to do with them. Henry continued to stand just inside the door, whimpering. For twenty minutes he stood there. Then he walked across the room and sat down on a chair with his back to the therapist. When he was told his time was up and asked if he would like to come back next week his answer was to get up, knock the chair over violently and kick the table. Then again he began to cower against the wall and whimper. The second time he came in—one week later—he went to the sand box and climbed in and began to dig a big hole with his hands—and, crouched down in the hole and tried to bury himself with the sand. He covered his face with his hands and stayed in the sand box for the entire period.

During the third interview he once again climbed into the sand box, but he threw some of the sand out on the floor—seemingly by accident. Then he got slowly out of the sandbox and went over to the table and tore up all the paper—crawled back in the sand box and rubbed his face against the sand. During the fourth interview he went around the room and upset all the toys in a feeble manner. Then he started to throw the sand wildly around the room. During the fifth interview he played at the sink with the water. He found the nursing bottle and filled it with water. He came over to the therapist, handed it to her and handed her the nipple. "What shall I do?" asked the therapist. "Fix it," he said—his first words to the therapist. She did. He sat down close beside her and drank from the bottle all the rest of the period.

During the sixth, seventh, and eighth interviews he drank from the bottle, seemed to be relaxed, even smiled wanly at the therapist. During the ninth interview he played with the doll—fed it with the nursing bottle—then suddenly tore the doll's clothes off, scratched its face—threw it all over the room. Then just as suddenly he sat down at the table, got out a piece of paper and some crayons and began to draw a house with many children playing outside. He looked at the therapist and said "This is us." He turned the paper over and drew another picture. He showed it to the therapist. "This is an all day sucker somebody gave me," he said. Then he drew a picture of a house and showed that to the therapist. "It's an empty house," he said. "Nobody in it." During the tenth interview he painted and talked to the therapist about his pictures—one "a storm with lots of black," one "a rainbow." During the eleventh and twelfth interviews he fingerpainted, talked a little to the therapist. At the conclusion of the twelfth interview his individual sessions were dropped and he was placed in a group of four boys for group play therapy.

Harold's card read as follows: "Extremely aggressive and unreasonable. Will not have a thing to do with another child. Defies the adults, but when he is convinced we are not going to fool with him he cringes and cries. But he is sneaky and can't be trusted."

Harold's behavior was very passive in the play room. He sat in the sand box with his back to the therapist for six interviews. During the seventh interview he dug holes and buried things in the sand. He selected the peg-pounding set and hammered on it until it split. Then he began to throw the sand around the room. Suddenly he stopped and went over to the doll and tried to take off its head. Not being able to do it he got the hammer and beat the doll—a rubber one. Then still not being able to break it he deliberately dug a hole in the sand box and buried the doll. Then he jumped on the place where he had buried it—then dumped the blocks on top of "the grave."

During the ninth interview he played with the nursing bottle—drank from it like a baby, crawled on the floor, found the toy gun and alternately "shot" the gun and "played baby." During the tenth interview he played with clay, made a clay dagger. He said very little to the therapist—and quite often kept his back to her. This was his last interview. He was stricken with polio—hospitalized—and received no further play therapy. In the hospital several months later he was given a Stanford-Binet. Received a score of 106.

Nancy's card read as follows: "Nancy is completely apart from the world of

reality. She does not talk—does not play —does not do a thing but stand around all day."

When Nancy came into the play room she looked like a little lost sheep. She sat on a chair and folded her hands in her lap. When the therapist explained the situation to her she looked blankly at the therapist, but made no other response. She sat there all during the first session. During the second interview she repeated the same behavior until the last five minutes. Then she reached over for the nursing bottle and sucked on that. During the third interview she started with the nursing bottle and clung to it all that session, and during the fourth, fifth, sixth, and seventh interviews. She smiled a little at the therapist, but did not talk. During the eighth interview she played with the doll, put it in the high chair, fed it, hugged it, very tenderly put it to bed. During the ninth interview she was drinking from the bottle again. During the tenth interview she divided her time equally between the doll and the nursing bottle. She placed the doll in the therapist's lap at the end of the hour and whispered "You take care of baby." During the eleventh interview she spent most of the time cooking a meal for the baby. Then she fed the baby and again put it to bed. During the twelfth interview she "played house" —This was her last interview. Suddenly the mother withdrew the child from the home. She was kept away for two months and then just as suddenly returned. It was not possible to resume therapy. School had started. The intelligence test given at school gave her a score of 100, but the teacher's comment on the record card stated—"a poor little deserted child—like a lost soul. No friends. No happiness."

However, schools and orphanages seldom have therapeutic opportunities available to help such children.

Eleanor's card read as follows: "Eleanor is the thorn in our side—stumbling into all sorts of trouble. She is almost blind— which I suppose causes lots of her trouble, but it doesn't make her any easier to live with. She is very very shy and does not speak even when she is spoken to. She does everything without purpose. Her trouble is that she knocks over things that belong to other children and breaks them—upsets paint, accidentally tears books and papers because of her fumbling ways. She will walk thru a game and not know she is ruining it or jar a puzzle someone is putting together and ruin it and not know she did it—or pick up some part of another child's toy and walk off with it and put it down some place else."

When Eleanor first came into the play room she stood in the middle of the floor, looking around slowly and then walked over to the doll house and fingered the furniture slowly, dropped it on the floor, shuffled around "looking for it," stepped on some of the pieces and broke them, looked fearfully at the therapist and not seeing any evidence of disapproval for this behavior continued to examine the dolls, dropped them, "accidentally" stepped on them. The behavior was slow and cautious with many furtive glances in the direction of the therapist. Then she walked over to the sand box knelt down beside it, picked up a little cup and a spoon and carefully filled the cup and "accidentally" spilled the sand on the floor. She repeated the frequent glances at the therapist, spilled more sand on the floor. Then she walked over to the table and looked at the drawing paper on the table, sorted thru it, held it up close to her eyes and "accidentally" tore it, another quick glance toward the therapist, and she moved on to the doll. She examined the doll carefully, hugged it, looked closely at its head and managed to twist it in such a way that the head came off. Then she went back to the sand box and poured sand into the cup so it spilled over on to the floor until the period was up.

During the second, third, fourth, fifth, sixth, and seventh interview this same kind of behavior was repeated—the only differences being in the selection of the material that was broken and a lessening of the number of glances at the therapist. She never spoke to the therapist—and when the therapist made comments she seemed to ignore them. When she came into the play room during the eighth interview she walked with more purpose, went over to the dollhouse and without any pretense swept the furniture out onto the floor and stamped on it. Then she went over to the sand box and picked up handfulls of sand and threw it on the floor. She looked defiantly at the therapist—"Watch!" she said defiantly—and threw sand out on the floor—with both hands. "You want me to be sure and see you throw it out on the floor?"

said the therapist with a questioning inflection. "Yes," Eleanor said. "I'll dirty it up." She ran over to the sink and got a jar of water and threw that on the floor. "How do you like that?" she asked angrily. Then she threw the jar down on the floor. She kicked it against the wall, trying to break it but was unable to and so went on to the rag doll. She twisted its head off again. (The rag doll's head, arms, and legs were sewed on loosely so they could be taken off by the children.) She pulled off its clothes and tore them, pulled off its arms and legs. Then she went over to the easel and poured the paints on the floor, pushed over the chairs, went over and kicked the door. When told the time was up she stuck out her tongue at the therapist, pushed her, went out stamping her feet.

During the ninth and tenth interview she repeated this kind of behavior but talked more to the therapist—said she "hated Jimmy," "hated Mary," "hated Don," and named many other victims of her hatred. During this tenth interview she spent part of her time painting at the easel—no form to the painting—just smears of black and red. During the eleventh interview she drew a picture at the table with the crayons. She told the therapist she was going to be adopted and would never "come back here" again. During the twelfth interview she said this was to be her last time coming over. She was calmer than she had been at any time previously. She talked to the therapist about her "adoption" and said she was glad she was leaving this "dump." She spent her time drawing with the crayons—interesting designs that varied from very simple balanced designs to complex balanced designs and intermingled among the designs bits of chaotic disorganized scribblings. She was given an intelligence test at the conclusion of this period. Eleanor was placed on trial in the home of a middle-aged minister and his wife. She was returned to the orphanage at the end of six weeks because she had not attained the approval of her intended foster parents. It was not possible to resume therapy with Eleanor when she returned to the home. According to a follow-up report she remained a problem to the rest of them, "perhaps even worse than before," according to the matron's report "because now she *deliberately* does things to annoy us, says she does and says she doesn't care."

The behavior and emotional expression of these five children presents material that is comparable to the behavior of the children in Group I and Group II. Because of the inclusion of more verbalization that some of the children in this group used we see, perhaps, a clearer picture of the nature of the emotional life of these children and the influence that it had upon their total behavior.

In summing up, we see how these fifteen children used the series of play therapy experiences to state their Self—and note the skill with which they did it. In some cases we see the expected behavior replaced by aggressive behavior. We see the "deliberate" behavior that in some cases replaced the "accidental" behavior. We see the emotional expression in action over an extended period of time and note the periods of suppression that flared out into explosive, conflicted, unleashed activity. We see it change from the dangerous calm to the destructive, aggressive negative release of tensions. We see the intrusion of ambivalence in the expression—the variation between all negative or all positive attitudes to include some negative and some positive attitudes. We see destructive behavior gradually or suddenly replaced by constructive behavior. We see the symbolic behavior become personal behavior—where the child exchanges the toy-relationships for a person-to-person relationship.

If we analyze the play according to categories we might see in these fifteen summaries the kinds of play that judges could validate—kinds of play that might indicate a projection of personality adjustment in process.

We might categorize the play as follows:—"initiating play," "release play," "symbolic play," "social play," "person-to-person play". Perhaps we could accept the definition of *play* as *"freedom or room to act"* rather than the usual use of the term when applied to children. These cases indicate that in many instances this experience was a grim

business—and not a recreational experience for these children. Perhaps further research might investigate the nature of each type of play as listed above—to determine, if possible, the relationship between the kind of play and its component in the process of adjustment. Perhaps we could test the following hypothesis:—*The child's play is a projection of his personality adjustment process in action—and his play represents the symbolic pattern of that organization changing from organization to disorganization to re-organization.* In the light of this thought we might ask: Is the "destructive" play of the child evidence of the emergence of constructive, reorganizing forces? Is it the symbolic movement of the disruption of the old pattern so that it can be changed into a more functional organization?

The play of these children challenges our thinking—and protests the use of simple labeling of behavior into compartments. The symbolic expression of the child who, after piling all the furniture in the middle of the room as a barrier between himself and the therapist, later seems to reduce it to a size that can give him objectivity when he piles the doll house furniture in the middle of the doll house room. And when Junior suddenly asks the therapist "What is your name?" and adds "*I* am Junior"—we try to seek an answer to the question posed by the emergence of the Self with such strength. What were the dynamics that

As we look at all these cases and at brought this about?

the sequence of their play patterns in each case it seems that the therapeutic experience is a present, cumulative, growing experience for the child that enables him to reorganize his perception of himself and to work thru his attitudes and feelings until he is able to establish a relationship of one kind or another with another person—the therapist who might be symbolic to the child of the adult or "other-person" world outside himself.

If the play therapy experience were ever demonstrated to be projection of the personality adjustment in process, then further studies could shed additional light on the place of the individual's capacity in the total Self.

We might then reexamine the question posed at the beginning of this paper —mental deficiency — symptom or disease?—and push beyond the redefinition to achieve more adequate educational procedures for the individuals so that the elements of emotional tensions, frustrations, conflicts, and deficiencies would not anchor the individual forever in one spot but would free him to develop more fully the capacities that are his. We would attempt to provide the kind of experiences that would give the child the strength to state his Self openly and freely so that he could obtain a healthy adjustment—so that each and every child could achieve the strength to look up and say "*I* am Junior. I like what I do."

Received January 10, 1949.

PSYCHOLOGICAL TESTS IN CLINICAL RESEARCH

By ROY SCHAFER
AUSTEN RIGGS FOUNDATION, STOCKBRIDGE, MASSACHUSETTS

PSYCHOLOGICAL tests are becoming increasingly important in clinical research. They have been applied to a wide variety of cases in two main ways: (1) To establish differential diagnostic indicators; for example, by seeking differences between schizophrenics and normal persons in their scatter of subtest scores on the Wechsler-Bellevue scale. That application is aimed at increasing the immediate usefulness of tests in the clinical setting, and the clinical basis of group division is the validating criterion. (2) To clarify and delineate personality differences between groups which have been set up on the basis of a gross clinical criterion; for example, by using the Rorschach test to define the personality differences between chronic alcoholics and normal persons. This application is aimed at increasing our understanding of the intrapsychic conditions under which psychopathological phenomena make an appearance. The validity of diagnostic personality indicators is taken for granted, and the clinical basis of group division is, in a sense, the variable to be validated against these indicators; that is to say, it remains to be seen whether the clinical division is a meaningful one in terms of personality dynamics. Both types of investigation appear to have increased in number and prestige far more rapidly than they have gained in methodological clarity.

This paper will be devoted to making explicit a number of assumptions that usually remain implicit in the use of psychological tests in clinical research. It will present an examination of these implicit assumptions for their consistency with the theoretical declarations of the researchers involved, derive some of the consequences of these assumptions for experimental design and interpretation of results, and, finally, formulate a general research orientation based on the results of the foregoing analysis.

NOSOLOGICAL AND STATISTICAL ASSUMPTIONS

1. *A diagnostic indicator is directly related to a diagnostic label.* This assumption is implicit in many research publications. Few would argue against the proposition that a diagnostic label is an abstraction and that, as such, it serves as a summary statement about the presence of a constellation of concrete characteristics which we may call *identifying characteristics.* The identifying characteristics need not be identical in all persons sharing any one diagnosis. The abstraction exists in the first place because it has this flexible quality.

Yet, most publications of efforts to establish differential diagnostic indicators in a psychological test pay little heed to the above obvious consideration. These studies relate diagnostic indicators directly to diagnostic labels and forget that the indicators must necessarily be expressions or reflections of identifying characteristics. It is not, for example, *schizophrenia* that should be correlated with an increase of pure color responses in the Rorschach test; it is *inappropri-*

ate affect or *the relative disappearance of the ability to interpose delay and thought between impulse and action.* Two schizophrenics might well differ greatly in their degree of detectable inappropriateness of affect and there would be good reason to expect their distribution of color responses in the Rorschach test to differ accordingly. These differences can hardly be established systematically so long as all schizophrenics are herded together into one group while individual differences in identifying characteristics are ignored. Similarly, not all obsessives manifest an equal degree of doubting or meticulousness, not all hysterics manifest an equal degree of emotional lability or "repressedness," and not all depressives are equally guilt-ridden or retarded. To combine all "neurotics" into one group for purposes of research into neurosis is obviously meaningless and fruitless.

What is required of clinical research using tests is that it concern itself more with the degree of presence or absence of specific identifying characteristics and less with traditional, psychiatric diagnostic labels. Because it brings us closer to the concrete case and yet does not preclude group comparisons, this shift of emphasis promises especially instructive results. Furthermore, this point of view enables us to understand and accept the absence of infallible diagnostic indicators: if the specific identifying characteristic need not always be present, there is little reason to expect its test "sign" to be invariably present. If, in addition, it is recognized that a specific identifying characteristic is not the exclusive property of any one group and that *constellations* of characteristics establish diagnoses, then it also becomes more meaningful that there is overlap between diagnostic groups with respect to the "signs" they show in their test records.

2. *A statistically established group trend characterizes all the members of the group.* Few would openly defend this assumption, yet there are numerous instances of its implicit acceptance in the "interpretation of results" sections of published articles. We find such statements as "alcoholics tend to. . ." or "schizophrenics are. . ." when actually the statistics have merely established that *relatively many* schizophrenics or alcoholics show a certain trend. The question still remains, "What about the remaining members of the group—those who do not show the trend?" It is arbitrary to assume that a statistically significant pattern captures the *essence* of group-membership.

A related erroneous assumption is seen when a mild statistical trend is assumed to indicate that all members of the group show the trend mildly. Too often the truth of the matter is that a small sample has been used, a few extreme score-deviations have occurred among the members of one group, and as a result, the test of significance of the difference is misleading. The literature is replete with contradictory findings due to this sampling problem.

This is not to say that the few great score-deviations are necessarily meaningless; it may well be in many instances that a relatively infrequent but nevertheless important identifying characteristic is present in the clinical picture of the few subjects with highly deviant scores. This point of view may be expressed in such a statement as, "If anyone gives a contamination response in the Rorschach test, it is likely to be a schizophrenic who does so, but not many schizophrenics give contamination responses." In other words it should not necessarily be assumed that a statistically insignificant difference implies that there is no difference between groups. This is a third and related type of as-

sumption about the statistical results of group comparisons that can confuse our understanding of our findings.

3. *The use of a large sample is always desirable since a large sample will cancel out individual differences that might otherwise becloud trends in our data.* This assumption appears to miss the point of clinical research entirely. Individual differences are to be established and understood and not cancelled out. Suppose that a study of the Rorschach test records of schizophrenics is to be undertaken. The proponents of this assumption would argue that a large sample is necessary since some of the subjects may be depressed at the time of testing, others may be excited, others may be in a panic, or deteriorated and bland, etc. In a large sample these "irrelevant" variations would cancel each other out to a large extent and we would then get a clear idea of what the *typical* or *essential* schizophrenic record is like. From the point of view advocated in this paper, the reply to this assertion would be, "But there is no typical schizophrenic record! There are only the records of depressed, excited, panicky, bland, etc., schizophrenics. You are reifying the abstraction *schizophrenia* and thereby violating the actual nature of our clinical material. If you say this will leave us with samples that are too small to allow for solid statistical analysis, we can only reply that the shaky results from a small sample have greater heuristic and clinical potential than the formidable summary scores of huge but inherently chaotic samples."

ASSUMPTIONS ABOUT TEST RESULTS

1. *A score is an adequate micro-unit for research.* Today, all but a few clinical psychologists would insist that the psychological significance of any one score depends on the context of scores and verbalizations in which it occurs. The recognition of this basic principle of test interpretation can be attributed to the inherent logical demands of our concrete clinical material buttressed by the theoretical expositions of gestalt psychologists. Yet, it appears that relatively few psychologists have seen fit to work out and apply the methodological implications of this principle for research using psychological tests.

The majority of publications of test research takes for granted that traditional statistical techniques are adequate for the gathered test data. The techniques referred to center around establishing the significance of the difference between single score-averages (t-test) and the significance of the difference between single score-distributions (chi-square test). In both instances the single *score* is treated as the fundamental unit of the test results. Thus, there are many published studies comparing the average incidence of each of the Rorschach test scores in two or more groups. Significant differences between averages are then held to indicate significant and *specific* personality differences between the members of the groups involved. This procedure therefore implicitly assumes that a score retains the same significance regardless of the context in which it occurs. Methodological confusion is further increased when the researcher tries to compensate for his initial disregard of individual score-patterns by interpreting the pattern of obtained averages. Thus the Rorschach test scores of a group may be averaged and the table of averages then interpreted as if it represented the individual scores of the "typical" group member. Furthermore, how is one to establish the specific shading of meaning of a group average if not by the circular method of borrowing from current clinical theory about the personalities of the members of the group? The

faulty logic of this approach is so striking that it is difficult to see how it can be implicit in so many publications.

The alternative to considering the score the adequate, meaningful micro-unit is considering the interpretation the meaningful unit. This approach recognizes that the same score can have different meanings and that the same meaning can be conveyed by different scores. It also recognizes the fact that scores in general do not adequately convey all the implications of a response or set of responses. Such factors as style and level of verbalization and conceptualization are as a rule altogether ignored by scoring units. In the Rorschach test, the response, "A lovely red and orange butterfly with delicate wings!" differs significantly from, "A butterfly. (What made it look like that?) The shape: wings and a body. (Anything else?) I suppose the color had something to do with it." Yet both responses on Card VIII would be scored *D FC* + *A*. A competent interpreter would not be likely in an individual case to consider these responses identical in meaning just because their scores are identical. To say that both responses indicate adaptive efforts places the concept "adaptive efforts" on so abstract a level that its clinical usefulness becomes questionable.[1]

Why then should so many competent workers abuse this principle in their studies of groups? One usual answer to this quesition is that the use of scores paves the way for a closer approach to "objectivity." The other side of this answer is that interpretation of verbalization is "too subjective." This answer has two weaknesses. First it ignores the fact that in individual cases interpretation of verbalization is a routine and indispensable part of the analysis of results. It therefore implies that psychologists may use a double standard in their work: one set of scientific principles for private practice and one set for public (academic) approval. This is a position few of us would like to be caught in. The second weakness of the answer that appeals to "objectivity" lies in its assumption of what makes for objectivity. Objectivity is bound first of all to respect the concrete clinical material at hand. If our scores fail us in certain instances by overabstracting the concrete material, then we should handle the scores flexibly, just as we would handle any tool of understanding. Scores are not meant to replace psychological thinking; they are designed to facilitate it, and as such they can be relegated to the background when this is warranted by the logic of the problem.[2]

Another usual answer to the question why scores are treated as adequate units in research is that they facilitate inter-individual and intra-individual compari-

[1]It may be supposed that the more closely two groups are equated, particularly if subcultural and personality factors are taken into account, the more likely it is that single score comparisons will yield meaningful results. This expectation is based on the proposition that fundamental personality dispositions define and limit the modes of expression of various psychological experiences, and that to the extent that two persons possess fundamental dispositions in common, they will express their psychological experiences in common forms. Thus, for example, highly intelligent, obsessive characters with lower middle-class, urban backgrounds are likely to resemble one another closely in the way they experience, express, and defend themselves against anxiety. As a group they would probably differ greatly in this respect from, for example, naive, repressive persons with poor rural backgrounds. The concept "polarized samples" may be used to convey this proposition. Empirical verification of the proposition is, of course, necessary and should be instructive.

[2]It may be argued that this difficulty can be overcome by devising additional and more subtle scores. It is difficult to conceive how these new scores would make any real difference, however, unless they took contexts and psychological implications of verbalization into account, thereby conforming in spirit to the position advanced in this paper.

sons. This answer holds water only if it can be simultaneously asserted that this procedure does not abuse or ignore the manifold implications of the concrete material, a problem that has been discussed in the preceding paragraphs. Questionnaires have always been popular in research because their results are so statistically pliable and lend themselves so nicely to comparisons, but in the main questionnaires have been of little real use in the clinic because by their very structure ("Answer Yes, No or Maybe.") they ignore the complexity of human responses. A more or less vague comparison of relatively precisely established facts is heuristically superior to a more or less exact comparison of relatively vaguely established facts, as Koehler has argued in his book *Gestalt Psychology* [1].

The third defense of score-oriented research and the one most worthy of consideration is that we are involved in a methodological dilemma. To state the dilemma: on the one hand we abuse our material by treating it as a collection of single scores; on the other hand we cannot effectively validate our interpretations without making comparisons and we cannot effectively make comparisons without using scores. We can't live with scores and we can't live without them.

This, however, is only an apparent dilemma. Its solution follows from an acceptance of, on the one hand, the principle that the clinical correlates of test patterns are identifying characteristics in syndromes and not diagnostic labels, and, on the other hand, the principle that interpretations rather than scores are the proper units for research. An interpretation may be defined as a prediction from test results that a certain pattern of behavior, thought and feeling will characterize the patient in his everyday life. This prediction can be checked against careful clinical personality examination. If, and only if, the principles of interpretation used are set forth in detail, with thorough elaboration of contexts which indicate this or that personality characteristic — if the cloud of intuition is lifted from the interpretive procedure — then the validity of our procedures can be established.

It can, for example, be stated that certain scatter patterns in the Bellevue scale, certain modes of concept formation in the Sorting Test, certain modes of approach to story-telling in the Thematic Apperception Test, as well as a general quality of verbalization throughout the tests, all point to a rigid reliance on obsessive-compulsive defenses; similar statements can be made with regard to repressive defense, control of passive needs, stability of object attachments, acceptance of a mature heterosexual role, and the like [2, 3]. These statements can be made on the basis of accumulated clinical experience. They can be put to the test by doing "blind" analyses of a mixed series of cases and attempting to single out the most obsessive, the most passive, etc. If this validational test yields positive results repeatedly, then the interpretive contexts validated thereby can be employed for studies where personality differences are to be established between members of two groups set up on the basis of some gross clinical criterion, such as alcoholism or anti-semitism.

In this later step of research it is again required that interpretations be made "blind"; that is, without knowledge of the group membership of each subject or patient. Qualitative analysis is too pliable by preconceptions. Thus, in a study of alcoholics, a mixed series of normal and alcoholic cases might be interpreted "blind" with reference to particular variables, such as low anxiety tolerance or control of passive needs. After the individual analyses are com-

pleted, the group membership of the subjects might be disclosed, the analyses sorted into the appropriate subgroups, and comparisons of interpretations drawn. If the interpretations differ systematically from one subgroup to the next, the research has been successful. This general approach paves the way for reasonable certainty of interpretations in individual cases in a manner that the group-average, score-manipulation approach can never do. This recommended approach might be called *controlled clinical experience*; its uncontrolled counterpart is the "clinical experience" that all would agree is a prerequisite of efficient and accurate use of tests in a clinical setting. Score-oriented research is not ruled out by this analysis, but it must be deferred until the qualitative analyses are completed; otherwise the qualitative analyses, if carried out by the same investigator, will be prejudiced by knowledge of the group-membership of the single cases.

2. *A single test is an adequate macro-unit for research.* This assumption is evident in the almost complete absence in the literature of group studies using more than one test. Most of all, the published studies have used an intelligence test or the Rorschach test. As a rule, only single case studies have been presented to demonstrate the usefulness of a battery of tests. The use of one test in research, as well as in clinical practice, brings with it certain almost inevitable consequences for the understanding of the test, and of personality and maladjustment. In the case of the Rorschach test this is most evident, since attempts have been made to build a theory of the "total personality" around the major personality trends indicated by this test. Simultaneously attempts have been made again and again to extend the interpretive potentialities of this test. The Rorschach test is currently much in vogue as an x-ray of "the unconscious"—an extension that is inevitably pretentious if rigidly adhered to. It is also used to "establish" significant historical sequences in the subject's psychosexual development. The thoughtful psychoanalyst generally bides his time over many months of intensive scrutiny of highly-charged, directly personal material produced by his patient before attempting such reconstructions. The line between Rorschach test interpretation and "wild psychoanalysis" is often difficult to discern in current publications.

This particular effort to stretch the usefulness of the test to the absolute limit appears to have led to, or perhaps to have been based upon, a failure to draw a clear line between test interpretation and clinical speculation about test findings. If a patient is found to be compliant through his Rorschach test responses, it may then be possible to speculate about patterns of impulse and defense which underlie this compliance, but this is no longer a Rorschach test interpretation. The accuracy or inaccuracy of these speculations can bring neither credit nor discredit to the test but only to the clinical astuteness of the interpreter. Unfortunately speculation of this type on the basis of the relatively incomplete personality picture which can be derived from this test or other tests cannot but lead to mechanical application of current, apparent clinical truths and thereby block the way to further insights into personality dynamics.

Tests vary from case to case in the depth and breadth of their instructiveness, and different tests vary in their instructiveness in any single case. It therefore seems advisable to consider a *battery of tests* as the basic macro-unit. The flexible way of approaching the "total personality" is to approach its

various aspects directly with a variety of appropriate instruments, not to make inferences about inferences and thereby spin elaborate fantasies around relatively few observations or findings.

Because the patterning of intelligence functions throws considerable light on the current status of ego organization, an intelligence test is important in a battery (preferably the Wechsler-Bellevue scale with adults). Because analysis of semicontrolled fantasies facilitates insight into prevailing needs and defenses against needs, as well as aspirations and evaluation of the self and other important figures, the Thematic Apperception Test is valuable. And because of the many advantages of the Rorschach test, which need not be summarized here, it is indispensible in a battery of tests. These three tests, supplemented perhaps by a word association test and a test of concept formation, yield a view of personality which surpasses in depth, breadth, specificity, and accuracy that which can be derived from the use of any single test.

SUMMARY

Research that proceeds by way of identifying personality characteristics rather than diagnoses as clinical criteria, individual variation as a focus of interest along with group trends, "blind" qualitative analysis prior to statistical comparison, interpretations as micro-units of analysis and batteries of tests as macro-units of investigation, appears to hold great promise for personality investigation.

Received February 24, 1949.

REFERENCES

1. KOEHLER, W. *Gestalt psychology.* New York: Liveright, 1929.
2. RAPAPORT, D., GILL, M. AND SCHAFER, R. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945, Vol. I; 1946, Vol. II.
3. SCHAFER, R. *The clinical application of psychological tests.* New York: International Universities Press, 1948.

A FACTOR ANALYSIS OF DISCRETE RESPONSES TO THE RORSCHACH INK BLOTS

By J. R. WITTENBORN
YALE UNIVERSITY

IN scoring a Rorschach protocol, it is customary to classify each response with respect to several classes of factors, such as content, portion of ink blot interpreted, and ink blot characteristics determining the response. There are standard categories within each of the classes of factors; when a response is scored, the content category in which it falls, as well as its location category and its determinant category, is designated by the scorer. As a second step in scoring, the total number of responses falling in each category is ascertained. Regardless of the precise manner in which the categories are used, the device of categorizing the response and then finding the total number of responses in each category carries two implications:

1. All of the responses falling in a given category are similar in some behavioral respect.

2. The psychological significance of responses falling in a given category is different in some respect from responses placed in other categories.

In an earlier study [10], statistical tests of these assumptions were based on analyses of responses to the Harrower-Erickson check list [5]. The analyses afforded negligible support for the assumptions. Certain of the findings are specifically relevant to the present investigation:

1. The portion of significant relationships among pairs of check list responses was greater than would be ascribed to the operation of chance alone.

2. The tendency for significant relationships between pairs of movement and color responses to be less predominant than significant relationships between pairs of movement responses or between pairs of color responses was not sufficiently marked to support the implications of Assumptions 1 and 2.

3. When pairs of responses were classified on the basis of the number of scoring categories (location, determinant or content) they had in common, it was found that the correlations between pairs of responses having two or three common scoring categories were significantly greater than correlations between pairs of responses having one or no common scoring category. Although it was statistically significant, the trend was not marked and there were numerous important exceptions.

As a result of these findings it was decided to submit the discrete check list responses to a statistical analysis which could provide an answer to the following questions:

1. The intercorrelations among the responses could not be satisfactorily predicted by certain deductions from Assumptions 1 and 2; is there some structure, some systematic clustering effect among the check list responses or are their relationships haphazard and unpatterned?

2. If there is a discernable pattern of relationships among the check list responses, will it be congruent with any current practice or belief concerning the Rorschach?

3. In the earlier study, only the movement and color responses were submitted to a specific examination. Can it be readily demonstrated that other determinants (e.g., form or animal movement) or location considerations (e.g., the distinction between whole responses and detail responses) play a discernable role in determining the inter-relationships among the responses?

4. In the check list, Harrower-Erickson included a number of responses which are generally taken to have a pathological implication; are healthy responses more highly correlated with each other than they are with unhealthy responses and will the correlations show that the unhealthy responses have more in common with each other than they have with the healthy responses?

Factor analysis of the intercorrelations of a carefully selected sample of check list responses is a possible means of providing answers to these questions.

THE DATA

The data for the present analysis were the responses of 247 Yale undergraduates to the Harrower-Erickson check list. The check list is attractive because it makes feasible a study which must employ discrete responses. There are, however, special considerations which must be observed in the evaluation of analyses based upon check list data:

1. Using check list data may yield results which are somewhat obscured by the suggestive effect of presenting the subject with a series of possible responses. Because of this possibility trends revealed by check list data may not equal the strength of trends inherent in other types of Rorschach responses. Nevertheless, the *pattern* of any trends inherent in other types of Rorschach responses should be discernable in trends among check list responses.

2. The check list procedure excludes many useful informal features of the individual Rorschach. In definitive analyses this constriction may be desirable because the check list responses are not obscured by variable features of the examiner's procedure, e.g., coaxing, expectant pauses, inquiring glances or differential emphasis in the inquiry.

The check list responses selected and their tetrachoric intercorrelations are shown in Table I. The responses included were selected in order to provide answers to the questions raised.

THE ANALYSIS

The intercorrelations among eighteen responses were submitted to a centroid analysis and the centroid factor loadings are shown in Table II. After the sixth

TABLE I

TETRACHORIC INTERCORRELATIONS OF EIGHTEEN SELECTED MULTIPLE-CHOICE RORSCHACH RESPONSES

No.	Card	Response	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1.	I	A bat																	
2.	I	Eyes glaring at me	.15																
3.	II	Little faces on the sides	.04	.29															
4.	II	Two bears rubbing noses	.00	−.01	.06														
5.	III	Two birds fighting	.06	.11	.45	.23													
6.	IV	Head of an animal	−.12	.20	.05	.10	.11												
7.	IV	A man sitting down	−.09	.24	.00	−.18	.00	−.27											
8.	IV	A man seen from below	.24	.21	.06	.10	.07	.02	.50										
9.	V	A ballet dancer	.10	.15	.07	.13	.16	.25	.20	.42									
10.	V	A bird flying	−.06	.00	.11	.00	−.03	.32	.03	.12	.23								
11.	VI	Two kings' heads with crowns	.07	.21	.07	.03	−.08	.10	.13	.36	.15	−.01							
12.	VI	Male and female organs	−.05	.06	−.10	−.05	.02	.18	.22	.12	.10	.10	.06						
13.	VII	Smoke	−.11	−.05	.41	−.03	.40	.16	.05	.00	.15	−.11	−.01	−.18					
14.	VIII	An orange and pink butterfly	.02	.05	.16	.15	.06	.21	.10	.19	.10	.01	.10	.10	.08				
15.	VIII	Fire and ice	.22	.11	.10	.10	.20	−.10	.26	.34	.16	−.17	.07	.30	.40	.20			
16.	IX	Two witches	.10	.01	.11	.17	.23	.28	.12	.22	.23	−.18	.32	.00	−.14	.45	.40		
17.	IX	A forest fire	.06	−.06	.10	−.04	−.05	.12	−.20	−.07	.44	.10	−.07	−.11	.05	.05	.20	.09	
18.	X	Two people	.13	.10	.22	.12	.36	−.19	.20	.40	.27	.10	.29	.05	.02	.10	.21	.33	.22

TABLE II

CENTROID FACTOR LOADINGS

Response	I	II	III	IV	V	VI
1.	+.16	+.15	+.08	—.27	—.11	+.12
2.	+.33	+.19	+.16	+.22	—.21	—.10
3.	+.41	+.25	—.33	+.26	—.22	+.18
4.	+.17	—.13	—.17	—.13	—.21	+.06
5.	+.43	+.18	—.17	+.21	—.17	+.22
6.	+.27	—.56	—.36	+.33	—.25	—.31
7.	+.29	+.33	+.42	+.10	+.24	—.32
8.	+.61	+.22	+.46	—.14	—.03	—.10
9.	+.59	—.29	+.23	+.12	+.27	+.26
10.	+.14	—.27	+.25	+.32	—.12	+.19
11.	+.35	+.08	+.25	—.18	—.20	—.22
12.	+.18	—.08	+.21	+.12	+.16	—.23
13.	+.24	+.29	—.54	+.34	+.28	—.05
14.	+.41	—.25	—.18	—.16	—.06	—.28
15.	+.54	+.17	—.17	—.26	+.50	—.24
16.	+.50	—.13	—.20	—.46	—.19	—.25
17.	+.21	—.27	—.07	—.08	+.27	+.28
18.	+.53	+.19	+.12	—.02	—.09	+.31

factor had been removed it was apparent that the variance of the distribution of residual correlations was less than the variance of a distribution of correlations based on random samples of 247 cases drawn from a population in which the true correlation was zero, and none of the residuals was significantly greater than zero.

The centroid factors were rotated orthogonally to a position described with respect to the response vectors by the rotated factor matrix shown in Table III. An inspection of Table III leads to

TABLE III

ROTATED FACTOR LOADINGS, ORTHOGONAL

Response	I	II	III	IV	V	VI
1.	—.03	+.30	+.12	—.06	+.19	—.08
2.	+.01	+.19	+.29	+.22	—.26	—.18
3.	+.28	+.43	+.05	—.09	—.46	—.08
4.	+.29	+.20	—.03	+.05	+.09	+.04
5.	+.17	+.44	+.11	—.02	—.37	+.01
6.	+.70	—.20	—.01	+.39	—.25	+.20
7.	—.30	—.11	+.62	+.01	—.15	—.17
8.	—.14	+.32	+.71	+.15	+.05	—.06
9.	—.03	+.23	+.42	+.11	—.14	+.62
10.	—.09	+.11	—.00	+.41	—.19	+.29
11.	+.07	+.16	+.42	+.20	+.15	—.18
12.	—.05	—.19	+.33	+.13	—.08	+.08
13.	+.26	—.04	+.07	—.47	—.58	—.02
14.	+.50	—.01	+.31	+.06	+.10	+.09
15.	+.21	—.04	+.64	—.51	+.01	+.12
16.	+.57	+.20	+.38	—.02	+.33	—.05
17.	+.06	+.10	+.04	—.14	+.05	+.49
18.	—.03	+.56	+.30	—.01	—.13	+.10

the impression that an oblique solution might provide a better description of the relationships among the responses than any orthogonal arrangement of the factors. Accordingly, a series of oblique rotations were performed with the aim of maximizing the number of zero loadings. The factor loadings resulting from the oblique rotation are shown in Table IV.

TABLE IV

OBLIQUE FACTOR LOADINGS

Response	A	B	C	D	E	F	X
1.	.03	.29	.01	—.05	.04	—.09	.20
2.	.09	.14	.36	—.08	.23	—.12	—.32
3.	.02	.09	—.04	.00	.63	—.10	—.25
4.	.25	.22	—.15	—.11	.09	.05	.11
5.	—.01	.07	.00	—.04	.54	.01	—.25
6.	.61	—.03	—.05	—.14	.06	.30	—.30
7.	—.07	—.03	.65	.20	—.06	—.16	—.22
8.	.13	.33	.56	—.02	.04	—.03	—.09
9.	—.04	—.20	.09	—.03	.19	.62	—.27
10.	—.04	—.15	.06	—.25	.12	.37	—.41
11.	.30	.39	.38	—.08	—.09	—.13	.03
12.	.08	—.11	.31	.07	—.12	.11	—.18
13.	—.13	—.31	—.05	.38	.51	—.12	—.20
14.	.52	.23	.08	.03	—.07	.10	.11
15.	.12	.00	.20	.45	.04	.00	.23
16.	.63	.54	.05	—.00	—.09	—.05	.37
17.	—.08	—.18	—.25	.04	.07	.45	.06
18.	—.07	.19	.10	—.07	.40	.10	—.12

In the expectation that the rotations would be simplified, a reference vector was added thus resulting in seven oblique factors.

Although the basis for the response clusters revealed by the factor analysis appears to be indefinite, a discussion of the various clusters and the characteristics of the responses which determine them is possible.

Response Group A	Orthogonal Factor I	Oblique Factor A
Head of an animal (6)	.701	.615
An orange and pink butterfly (14)	.496	.520
Two witches (16)	.565	.629

These responses are alike only in that they are based upon large details of the ink blot.

Response Group B	Orthogonal Factor II	Oblique Factor E
Little faces on the sides (3)	.401	.634
Two birds fighting (5)	.436	.541
Smoke (13)	.040	.507
Two people (18)	.556	.400

The stable members in this cluster of responses (3, 5, and 18) have several common features: they are usually based on large details of the card, usually do not involve color but are drawn from colored cards, and in the case of responses 3 and 18 involve human content. Response 13 has a doubtful status with respect to this cluster inasmuch as it appears only in the oblique solution and, from a qualitative standpoint, is seemingly different from all of the others.

Response Group C	Orthogonal Factor III	Oblique Factor C
A man sitting down (7)	.620	.649
A man seen from below (8)	.705	.555
A ballet dancer (9)	.422	.090
Two kings' heads with crowns (11)	.421	.376
Fire and ice (15)	.637	.202

Responses 7 and 8 are drawn from the same card and are practically identical in content. They were placed in the analysis originally for several reasons, including the expectation that they might form the nucleus for a factor in which human movement was an important component; however, not all of the human movement responses are correlated with Factor III or C. The oblique solution indicates that responses 7 and 8 comprise a doublet and does not encourage discussion of Factor III as a meaningful cluster.

Response Group D	Orthogonal Factor IV	Oblique Factor D
A bird flying (10)	.410	—.252
Smoke (13)	—.462	.384
Fire and ice (15)	—.506	.451

A comparison of the loadings for these two factors suggests that responses 13 and 15 bear an important functional similarity to each other. These responses are usually based on the entire card, and the form of the ink blot appears to play little or no role in their determination.

Response Group E	Orthogonal Factor VI	Oblique Factor F
A ballet dancer (9)	.62	.62
A bird flying (10)	.29	.37
A forest fire (17)	.49	.44

These three responses are alike in that they usually involve some kind of movement and usually involve the whole ink blot. They are obviously different, however, with respect to abstract Rorschach scoring determinants. Response 9 is a human movement response. Response 10 is determined by animal movement and 17 is a poorly controlled color response which often involves inanimate movement.

Orthogonal Factor V was inverted in the oblique rotations and contributed to the formation of Factor E. With Orthogonal Factor III, Orthogonal Factor V contributed to the variance of Oblique Factors B and X. No satisfactory description can be offered for Orthogonal Factor V. Oblique Factor X is a bipolar factor comprising two tests, each of which involves movement (a bird flying and the two witches). Factor B is determined by responses 11 and 16 which are based upon large details of the ink blot and have a human content, often of a fantastic and mythological nature.

DISCUSSION

The results of both the orthogonal and the oblique rotations indicate that there is some systematic clustering effect among the check list responses. The pattern of relationships discernable in the present sample of responses does not

appear to be precisely congruent with current practices or beliefs concerning the Rorschach. There is strong indication, however, that location considerations may play as great a role as determinant factors in accounting for the interrrelationships among the responses.

Response Group D is of particular interest because of its relevance to the fourth question posed at the beginning of this paper. Response 13 (Smoke) and Response 15 (Fire and ice) have important negative projections on orthogonal Factor IV and also appear together on oblique Factor D which is derived from Orthogonal Factor IV and which bears a negative relationship to the other oblique factors. These responses, commonly accepted as having pathological implications, are not only correlated with each other but bear a similar pattern of correlations with the other responses. Thus it appears that these responses tend to distinguish between individuals in a similar fashion and offer evidence that relevance to pathology may be an important factor in the intercorrelation of at least some Rorschach responses. Nevertheless, other responses suggestive of pathology (17 and possibly 8 and 12) are not part of this cluster.

The results of analyses A and B [10], as well as the results of the present analysis, suggest that the patterns among Rorschach perceptions call for the introduction of new or considerably revised concepts. It is not implied that the use of Rorschach perceptions or responses by skilled workers in their psychodiagnoses is invalid. It seems very probable, however, that the explicit conceptualizations currently employed in teaching and discussion are either insufficient or misleading. It appears likely that advancements in the study of the relationships between perceptions (Rorschach responses) and personality attributes are more dependant on rigorous tests of refined or reformulated concepts[1] than on the accumulation of larger samples, on a compulsive refinement of gratuitous scoring procedures, or on a wider variety of merely "statistically significant" differences between vaguely defined groups.

It is evident from the analysis that the intercorrelations among the responses are not primarily a reflection of abstract scoring categories. This may possibly be regarded by some readers as challenging to the status of the Rorschach as a clinical instrument. It is suggested, however, that the present findings indicate that abstract Rorschach scoring categories are less important than commonly supposed. Familiarity with the Rorschach over a period of years has led the writer to surmise that the content of Rorschach responses (including the perceptual and associational processes they involve for the individual) has a relevance to fantasies, motivations, and adjustment patterns which is equal to, if not greater than the relevance of the greatly emphasized abstract scoring categories. In the present study, human movement responses are distributed among the factors in a manner which suggests that their content and other aspects not reflected by abstract scoring categories are important determiners of the patterns in which they are intercorrelated.

The exact nature and degree of the limitations in current Rorschach beliefs and practices can be stated with confidence only after the completion of further analyses based upon data gathered by other procedures, upon samples comprising different types of subjects, and upon experimental designs which test various interpretations of some of

[1]In some cases, new concepts must be developed by making explicit much of what is now implicit in Rorschach evaluations.

the implicit features of Rorschach practices. Because of its known practical limitations, a Rorschach check list is of limited interest per se; nevertheless, the form lends itself to quantitative analyses which in turn yield hypotheses relevant to and readily tested by the individual form of the test. With perhaps some exceptions it may reasonably be supposed that the responses which the individual has checked on the check list comprise a sample of the perceptions which may be elicited from him by the ink blots. Therefore, in the absence of specific contraindications, it is justifiable to generalize the results of certain statistical analyses (which are convenient only with data yielded by the check list form of the test) to relevant assumptions which have implications for all forms of the test. Consequently, inferences from the present study have been employed by the writer in the formulation of hypotheses which are being tested with data yielded by the individual Rorschach procedure.

SUMMARY

The present study comprises a factor analysis designed to answer certain specific questions. Since the study is concerned with a statistical evaluation of the interrelationships among discrete responses to the Rorschach ink blots, the responses were secured by means of the Harrower-Erickson check list. From the factor analysis, which involved 18 responses and was based upon 247 students, the following conclusions may be offered:

1. The discrete responses are intercorrelated in a fashion to yield factors some of which are relatively clear-cut from the quantitative standpoint.

2. An inspection of the factors yielded by both orthogonal and oblique methods of rotation does not offer evidence in support of some of the common beliefs concerning the Rorschach responses, particularly those which are reflected in abstract scoring practices.

3. There is evidence that certain so-called "pathological" responses possess an important functional similarity.

4. The data suggest that response content (with its implied associations and projections) may play an important role in determining the functional similarities and dissimilarities of certain responses, particularly those which involve movement.

Received December 20, 1948.

REFERENCES

1. BALINSKY, B. The multiple choice group Rorschach test as a means of screening applicants for jobs. *J. Psychol.*, 1945, 19, 203-208.
2. BECK, S. J. *Rorschach's test.* New York: Grune & Stratton, 1945, Vols. I and II.
3. CHESIRE, L., SAFFIR, M. AND THURSTONE, L. L. *Computing diagrams for the tetrachoric correlation coefficient.* Chicago: Univ. of Chicago Bookstore, 1938.
4. ENGLE, T. L. The use of the Harrower-Erickson multiple choice test in differentiating between well adjusted and maladjusted high school pupils. *J. educ. Psychol.*, 1945, 37, 550-556.
5. HARROWER-ERICKSON, MOLLIE R. AND STEINER, M. E. *Large scale Rorschach techniques.* Springfield, Ill.: Charles C Thomas, 1945.
6. KLOPFER, B. AND KELLEY, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1942.
7. RAPAPORT, D. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945. Vols. I and II.
8. RORSCHACH, H. *Psychodiagnostics.* Berne: Hans Huber, 1942.
9. THURSTONE, L. L. *Multiple factor analysis.* Chicago: Univ. of Chicago Press, 1947.
10. WITTENBORN, J. R. Statistical tests of certain Rorschach assumptions: Analyses of discrete responses: *J. consult Psychol.*, 1949, 13, 257-267.

THE RORSCHACH AS A MEASURE OF INTELLIGENCE

By W. D. ALTUS and GRACE M. THOMPSON
SANTA BARBARA COLLEGE
UNIVERSITY OF CALIFORNIA

RORSCHACH in his *Psychodiagnostics* [7] lists seven factors which he felt to characterize the intelligent subject. These seven consisted of good form perception, many *M* (human movement responses), a large number of whole answers, an emphasis on wholes or wholes and large details with small details not unduly overemphasized, orderly succession from whole to large detail to small detail, small percentage of animal answers, an optimum percentage of original answers. Unfortunately, it is difficult, if not impossible, to subject all of these seven attributes to experimental validation through quantitative means. For instance, it would be difficult to express in a number one's tendency toward what Rorschach called "orderly" succession; in addition, the scoring of "originals," shot through with subjectivity as it is, would make quantification thereof somewhat suspect.

The influential followers of the Rorschach method have mostly remained content to play exegete to his *Psychodiagnostics*, adding little and subtracting nothing from his list. Beck [2] holds that his scoring of the degree of organization in the Rorschach protocols (Beck calls this organization activity Z) serves as "an index of thinking power. Its essence is the capacity to grasp relations not perceived by others." Klopfer [5] claims that the number of content categories is related to intelligence. In the main, however, Rorschach's original observations regarding the work of the intellect in the interpretation of the ten ink blots appear to be as well accepted today as they were when Rorschach's basic work was written, almost a generation ago.

The present study represents an attempt to eke out, by quantitative means, certain correlates of intelligence in the group Rorschach, making use of all those "items" which lend themselves easily to numerical treatment. The population studied was the traditional elementary psychology student, in attendance at Santa Barbara College of the University of California. The first sample consisted of 128 students, who in 1946 were administered the Group Rorschach, following Munroe's directions [6]. The second sample consisted of 100 students in the summer of 1947. This second group took the group Rorschach twice, at the beginning and at the close of the six-weeks' summer session. These two administrations permitted the calculation of test-retest reliability coefficients for the Rorschach. Previous studies [4, 10, 11] had made use of the split-half, single administration of the ink-blot test in deriving of reliability coefficients for *M*, *W* and certain other factors. It was felt by the present writers that test-retest reliability would be a more realistic appraisal of items used rather than the occasionally spuriously high coefficients obtained by the split-half method.

As has been noted in previous publications [1, 3, 9], some seventy-five aspects of the Rorschach were quantified

and entered on cards for each individual student, thus permitting the usual interquartile item-analysis technique to be employed in the attempt to find correlates of intelligence. Two criteria of intelligence, both admittedly highly verbal, were made use of: (1) for the 1946 group, the Altus' *Measure of Verbal Aptitude;* (2) for the 1947 group, the *Ohio Psychological Examination, Form 21.* An inspection of the Q4-Q1 item analyses for the 1946 and 1947 groups revealed that 15 Rorschach items held, even though the criterion of intelligence was changed from one testing to the other. The 15 items, their tetrachoric validity coefficients, the cutting lines for the tetrachoric computation, and the reliability coefficient for each item all appear in Table I.

TABLE I

THE TETRACHORIC RELATIONSHIPS BETWEEN CERTAIN ASPECTS OF THE GROUP RORSCHACH AND TWO MEASURES OF VERBAL INTELLIGENCE, TOGETHER WITH THE TEST-RETEST RELIABILITIES OF THE RORSCHACH ITEMS

	Rorschach Item and Cutting Lines	Validity A* 1946	Validity B* 1947	Validity C* 1947	Reliability D* 1947
	N	128	100	100	100
1.	M, 4 and higher	.43	.34	.43	.93
2.	M/M+ FM, .60 and higher	.34	.12	.12	.58
3.	M/M+ Sum C, .50 and higher	.32	.22	.29	.72
4.	W, 5 or more	.13	.28	.28	.90
5.	W + Dd/W, 1.00-1.99	.17	.24	.16	.87
6.	Dd%, 0-14	.10	.19	.25	.65
7.	P, 7 or more	.22	.16	.03	.63
8.	Achromatic R/Total R 45% or more	.29	.22	.03	.40
9.	m, 1 or more	.27	.16	.13	.15
10.	FC/C + CF, 2.00 and higher	.20	.39	.12	.13
11.	F%, 0-59	.23	.31	.03	.67
12.	"Z"/R, 1.10 and up				
†13.	W%, 25 and higher	.16	.07	.00	.39
†14.	H/Hd, 3.00 and up	.20	.33	−.07	.77
†15.	FV + Y, 4 or more	.16	.13	−.07	.54

*The criterion for "A" is the Altus' **Measure of Verbal Aptitude**; for "B" and "C", the **Ohio Psychological Examination, Form 21.** "B" represents the validities for the first administration of the Rorschach and "C" the validities for the second administration. "D" gives the test-retest reliabilities of the Rorschach items.

†These last three items were not used in the Pearson product-moment coefficients of correlation noted in the body of the article.

It will be immediately noted that the most valid item of the 15 is the sheer incidence of human movement: When the cutting line is set at 0-3 versus 4 or more M, the resulting correlation with the *Measure of Verbal Aptitude* is .43; with the first Rorschach administration (1947) and the *Ohio Psychological,* the r is .34; the second Rorschach administration (1947) shows an r of .43 with the *Ohio.* The tetrachoric coefficient of reliability is also gratifyingly high, .93. The absolute number of Whole responses (0-4 versus 5 or more) shows consistent validity coefficients of .13, .28 and .28. The reliability of the number of Whole responses is also high, .90. The relation of Whole to Small Detail responses is also consistently valid, .17, .24, .16, and its reliability, .87, is quite adequate as well. The findings relative to these three aspects of the Rorschach test confirm the observations of the author of that test as quoted in the opening paragraph of this paper. The $F+\%$ as an indicator of intelligence does not hold for the groups here considered—though the coefficients were not calculated, an inspection of the item analyses revealed that they would be minimal or zero in character. It may well be, of course, that the relatively high level of intelligence of college students and the resulting homogeneity of their intellectual level vitiates the $F+\%$ indicator which is given precedence in Rorschach's ordinal listing of indicators of intelligence. In like manner, the $A\%$ was not considered because the item analyses showed that the results would have been equivocal, just as they were with the $F+\%$.

A further inspection of Table I reveals that three other movement responses show some validity: (1) the relation of human to animal movement (item 2) is fairly valid, though the re-

liability is not high, .58; (2) the relation of human kinaesthesias to the color sum (item 3) is relatively valid, .32, .22 and .29, and somewhat reliable, .72; (3) inanimate movement (item 9) shows consistent validity, .27, .16 and .13, but the test-retest reliability is low almost to the vanishing point, .15. It is obvious that the accuracy of scoring inanimate movement for the 1946 group must have been greater than for the 1947 group—for unless one makes such an assumption, it is impossible to account for its validity in the 1946 group being greater than the actual reliability, short, of course, of the intervention of chance factors. It is obvious from the statistics just given that the *m* component (for which we are indebted to Klopfer) would be a quite valuable addition to intelligence indicators in the Rorschach if a more accurate system of scoring it could be devised. It is also obvious that there is a significant tendency for the brighter college students to have more human than animal movement in their records and to have an *M*: *Sum C* ratio weighted on what Rorschach called the "introversive" side.

Item 6, which concerns the incidence of small detail visualized, shows that the brighter student tends to neglect small detail, especially in favor (item 5) of Whole responses. Rather interestingly, this same indicator — *Dd* per cent—was shown in a previous publication [1] to be significantly linked to the Sc scale of the *Minnesota Multiphasic Personality Inventory*. Two of the three validity coefficients for item 7 indicate that the brighter college student tends to "see" more populars (scored *à la* Beck) than the relatively dull college student. This tendency is not a reflection of the *R*, or total responses, since in the groups here considered the number of responses was not related to the criteria of intelligence. However, since the third validity coefficient for the number of populars visualized drops to .03, this Rorschach item may be held somewhat suspect. In like manner, the percentage of responses projected from the achromatic cards appears to be a valid indicator for the first two testings but the validity coefficient for the 1947 group on the second testing drops to .03. The percentage of achromatic answers appears also to be rather undependable since the coefficient of reliability is only .40.

The least reliable of all the items presented in Table I is the relation of *FC* to *C* and *CF*. The reliability coefficient of .13 for this item makes somewhat dubious the validity coefficients which precede it. How it is possible for the validity coefficient to have a larger value than the reliability coefficient in two instances is difficult to rationalize unless one posits that demon of the statistician, the "chance" factor. If it were not for the troublingly low reliability of this item, one would feel that the brighter college student was definitely characterized by the *FC* answer in contrast to the more "impulsive" and "primitive" *C* and *CF* ones.

Item 11 shows a quite understandable tendency for the college student with less pure form in his record to be the brighter individual. This finding is rather in the nature of an artifact, since one of the most significant factors in displacement of pure form — at the college level, at least—is the *M* component, already shown to be heavily saturated with intelligence. Item 12, which presents validation of Beck's *Z* component, is rather interesting: it is obviously a quite reliable factor (.87) but the first fine flush of validity shown, .40, drops quickly to .10 and .07 for the 1947 group. Item 13, which considers the Whole response as a ratio rather than an absolute number, has high re-

liability (.89) but its validity decreases to zero with the third coefficient. It might be reasonably inferred that the absolute number of Whole responses is more indicative of intelligence than the ratio of such responses to the total number given. The fourteenth item, the ratio of human detail responses to whole human responses, had fair validity, .20 and .33 for the first two applications of the Rorschach, but the third coefficient shows a slight negative value. The shading responses, depth and surface summed, gave promise, the first two administrations, of having some relation to intellectual factors; but the validity vanished on the third administration.

Thus far only individual items in the Rorschach have been considered in their relation to intelligence. A further exploration was attempted by combining "signs" of intelligence by giving a unit weight of one for each "sign" in the individual protocol. Only the first twelve of the fifteen items in Table I were thus employed. After each record was scored in this fashion, it was possible to compute Pearson product-moment coefficients of correlation with the two intelligence tests employed as criteria. The twelve-point total correlated with *Verbal Aptitude* .40; for the 1947 group, the first administration of the Rorschach yielded an r of .37 with the *Ohio*, while the second administration showed a marked regression in the validating coefficient, .24. The first two coefficients noted are significant at the .01 level, while the third drops to the .05 level of confidence. The Pearsonian reliability of the twelve-point set of items proved to be .71. Such a reliability is to be expected when items with so varying a degree of consistency (.13 to .93) enter into the gross score.

The validity coefficients for the set of 12 items, each bearing unit weight, were disappointingly low. It was conjectured that perhaps this apparent lack of validity was a function of the unreliability of certain of the 12 items, such as the scoring of color and of inanimate movement. For this reason it was decided to take the most valid and reliable item — which an inspection of Table I will show to be the M component—and to derive some Pearsonian validity coefficients by the use of it alone. The results were somewhat surprising: The absolute number of M correlated (Pearson product-moment) .37 with *Verbal Aptitude*; with *Ohio*, first administion of Rorschach, .41; with *Ohio*, second administration of Rorschach, .31. It is to be noted that all three of these coefficients are significant at the .01 level of confidence while only two of the r's for the laboriously computed 12 items were significant at this level. It would seem, then, that for college students in institutions with certain minimum entrance standards the sheer incidence of *M* is just as valid (and perhaps slightly more so) an indicator of intelligence as any possible combination of readily quantifiable items.

The Pearson product-moment coefficient of reliability for the *M* function proved to be .82. While this coefficient is .11 lower than the tetrachoric coefficient of .93 shown in Table 1, it is not unusual for tetrachoric coefficients to regress in this fashion. It will be remembered that the Pearsonian reliability of the 12 items was .71: the *M* function is not only more valid but is also more consistent than the set of items of which it is a part. It was further noted on the scattergrams where the *M* function was plotted against the criteria of intelligence that a curvilinear relationship seemed definitely to be present. The *Eta* between *M* and the *Verbal Aptitude* [9] proved to be .54, in comparison with the linear r of .37. The *Eta* for the first administration of the Rorschach in 1947

with the *Ohio Psychological Examination* proved to be .63. Linearity of regression does not hold in either instance; the *M* function, in its relation to intelligence at the college level, appears to be curvilinear. What seems to hold is this: While a sizable number of *M* (say six or more at the college level) is almost invariably evidence of superior intelligence, the absence of *M* cannot be interpreted as indicating inferior intelligence. It is true that the relatively dull have no *M* or only a few; but sometimes the relatively bright will show the same pattern. Why the *M* function is reduced in some bright people is not surely known; it may be that maladjustment plays a part in these cases; it may also be that high quantitative ability, in contrast with verbal, is deleterious so far as the projection of human kinaesthesias are concerned. Whatever the cause, the fact of the curvilinearity of regression of the *M* function on intelligence at the college level appears to be fairly well established.

Hertz [4] some fifteen years ago published a correlational matrix of certain Rorschach factors; included in the matrix was the IQ of her junior high-school student population. For this group, *M* correlated only .259 with IQ, about the same as the number of Whole responses (.241) and considerably under the percentage of originals, .398, and the percentage of *F*+, .460. One reason for the apparently lower saturation of *M* with intellectual factors in her study in comparison with the present one is that she used IQ's instead of raw scores or mental ages; the latter, hypothetically, should show higher correlational value for the wide gamut of chronological ages found in a junior high-school population. In the same study, Hertz reports a split-half reliability coefficient of .745 (N of 300) for *M* percentage. Vernon, also using the *M* percentage and the split-half technique, reported a reliability of .62 for 90 cases [11]. Thornton and Guilford [10] obtained split-half reliabilities for two groups of college students, *N*'s of 42 and 58: For the *N* of 42, the absolute number of *M* gave an *r* of .919; for the larger group, the *r* proved to be .768. It must be remembered that all of these reliabilities came from a single administration of the Rorschach. Whether such reliabilities have the same meaning as the test-retest reliabilities reported in the present study may be open to question.

An interesting *caveat* relative to the presumptive saturation of the *M* function with intellectual factors will now be entered. Sicha [8] in her Ph.D. research found that Negro adults of superior educational attainment and intellectual status had significantly fewer *M* responses than did a control group of Whites; in fact, the number of *M* found among the Negro group was slightly less than one-half that found among the Whites. Whether such a finding is true all along the continuum of Negro intelligence is, of course, unknown. If it were, it is clear that the extrapolation of intelligence from the absolute number of *M* in the Rorschach record would do an injustice to the Negro, especially if White norms were used.

Klopfer [5] has an observation concerning the Rorschach and Binet IQ's which may be of interest to record: "It is possible to evaluate a Rorschach record and to 'guess', in the majority of cases, within a range of ten points, what the intelligence of the subject in terms of a Binet IQ might be. . ." While Klopfer presents no quantitative buttressing of his general statement in his text, the writers feel that he might well better the higher of the two *Eta* values noted in this study, .63. A psychologist who resonates to the niceties of vocabu-

lary and diction could probably make a quite fair estimate of the intelligence implicit in a Rorschach protocol, even if he were totally ignorant of such subtleties as *M*, *Z*, *W*, *Dd* and *P*. If such a hypothetical psychologist were conversant with the scoring of the Rorschach, his considered judgment (not "guesses") should, in general, be quite valid.

SUMMARY AND CONCLUSIONS

A group of 128 college students were subjected to the group Rorschach in 1946 and the protocols were scored in the manner set forth by Munroe [6]. Another group of 100 college students were twice given the Rorschach in the summer of 1947. The first group took, as a measure of intelligence, the Altus' *Measure of Verbal Aptitude*; the second group took the *Ohio Psychological Examination, Form 21*. Seventy-five factors were extracted from each individual Rorschach protocol and the technique of interquartile analysis was employed to find items saturated with intelligence. Fifteen such items were abstracted; three validity coefficients and one reliability coefficient were calculated for each of the 15 items. Twelve of the 15 items were later lumped into a composite score; this score was validated against the two criteria of intelligence and a reliability coefficient for the score was also calculated. Finally, the absolute number of *M* in each protocol was validated against the criteria by the more rigorous Pearson product-moment technique of correlation; the reliability of the *M* function was obtained by the same technique.

The main findings follow:

1. The most valid and reliable indicators of intelligence in the Rorschach proved to be the absolute number of *M*'s and *W*'s, with the *M* being definitely the more valid and slightly more reliable.

2. The various factors considered (15 in number) varied widely in reliability from a low of .13 (tetrachoric) for the *FC/C+CF* ratio to a high of .93 for the absolute number of *M*.

3. It seems probable that the greater validity of *M* and *W*, in contrast with certain other Rorschach factors, is in part a function of their relatively quite high reliability.

4. Five of the seven Rorschach aspects which Rorschach claimed in the *Psychodiagnostics* [7] to be indicators of intelligence were checked in the present study. Three of the indicators, *W*, *M*, and the relation of *W* to *Dd*, proved to have definite validity, while two, *F+%* (Beck's form tables were used) and *A%*, did not appear to discriminate. The restriction in the range of intellect found among college populations may be responsible for the latter findings.

5. The *M* function by itself appears to be as valid an indicator of intelligence at the college level as any combination of Rorschach factors of the type extracted in the present study.

6. The relationship of *M* to intelligence appears to depart significantly from linearity. The *Eta* for *M* and one criterion of intelligence was .54; for the other criterion the *Eta* was .63.

7. A large number of *M* in an individual Rorschach protocol appears to be prima facie evidence of relatively superior intelligence; the converse does not, however, hold true since some individuals of quite adequate intelligence show few or no *M*.

8. The findings of this study should not be generalized beyond the population from which they are derived, that of students in a college which uses selective admission requirements.

Received January 10, 1949.

REFERENCES

1. ALTUS, W. D. Some correlates of the group Rorschach and the schizophrenia

scale of the group MMPI among two groups of "normal" college students. *J. consult. Psychol.*, 1948, 12, 375-378.

2. BECK, S. J. *Rorschach's test II. A variety of personality pictures.* New York: Grune & Stratton, 1945.
3. CLARK, J. H. Some MMPI correlates of color responses in the group Rorschach. *J. consult. Psychol.*, 1948, 12, 384-386.
4. HERTZ, MARGUERITE R. The reliability of the Rorschach ink-blot test. *P. appl. Psychol.*, 1934, 18, 461-477.
5. KLOPFER, B., AND KELLEY, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1942.
6. MUNROE, RUTH. *Prediction of the adjustment and academic performance of college students by a modification of the Rorschach method.* Stanford Univ. Press: Applied Psychology Monographs, No. 7, 1945.
7. RORSCHACH, H. *Psychodiagnostics.* Berne: Huber, 1942.
8. SICHA, M. H. *A study of the Rorschach 'Erlebnis-typus' of comparable White and Negro subjects.* (City and publisher not known.) 1939.
9. THOMPSON, G. M. MMPI correlates of certain movement responses in the group Rorschachs of two college samples. *J. consult. Psychol.*, 1948, 12, 379-383.
10. THORNTON, G. R. AND GUILFORD, J. P. The reliability and meaning of Erlebnis-typus scores in the Rorschach test. *J. abnorm. soc. Psychol.*, 1936, 31, 324-330.
11. VERNON, P. E. The Rorschach ink blot test, II. *Brit. J. med. Psychol.*, 1933, 13, Part III, 179-205.

VALIDATION OF THE ROTTER INCOMPLETE SENTENCES BLANK FOR COLLEGE SCREENING

BY JULIAN B. ROTTER, JANET E. RAFFERTY AND EVA SCHACHTITZ
OHIO STATE UNIVERSITY

THE use of the incomplete sentences technique for personality assessment was given decided impetus by its application in the armed services [2, 3, 7, 8, 9].[1] The version used by Rotter and Willerman [6] indicated relatively high validity for determining the degree of mental disturbance in patients in an AAF Convalescent Hospital, and their study demonstrated that it was possible to have both objective scoring and certain advantages of the projective techniques. The freedom of response, enabling the subject to express his own conflicts rather than conforming to a "yes", "no", "?" categorization, is retained, as well as a certain amount of ambiguity or disguise of purpose. The latter is perhaps better expressed by saying that the subject is less knowledgeable as to what is a "good" as against a "bad" answer. It seemed probable that the development of new scoring manuals could make this test useful in a variety of other screening problems.

The present study is concerned only with the validation of the test as used to obtain an overall score for degree of conflict or maladjustment of college men and women. Employed in this way the test might be given to incoming freshmen, or to special classes to determine which students are in need of personal help. In addition to such screening, the test results may provide significant material for individual clinical evaluation of the subject of particular value in structuring the first interview or counseling session. It has been applied in general experimental procedures in which adjustment and the effects of treatment processes were assessed, and in vocational and remedial counseling programs to determine the extent to which maladjustment is an important factor in the presenting problem.

[1]Reviews of previous work with incomplete sentences are given by Bell [1], Rohde [4, 5] and Symonds [9].

THE TEST AND ADMINISTRATION

The present test is a modification of the one used earlier by Rotter and Willerman [6], which was modified from forms used by Shor, Hutt, and Holzberg [7, 3, 2] at the Mason General Hospital.[2] The 40 items and instructions are presented in Figure 1.

No further instructions are given except to repeat the printed instructions and to urge subjects to complete all the items. Administration to a group of any number of subjects is possible. The approximate average time for administration is 20 minutes.

DEVELOPMENT OF THE SCORING MANUALS

Previous experience has suggested that a criterion group containing fewer cases each of whom has been carefully

[2]Rohde [5] has recently pointed out that some of the items of the incomplete sentence tests used in the Army are similar to items in a test copyrighted by her and Hildreth in 1941 and described in the psychological literature in 1946. The Rohde-Hildreth test was in turn modified from a test of Payne's which had not been described in the psychological literature prior to 1946.

COMPLETE THESE SENTENCES TO EXPRESS YOUR REAL FEELING. TRY TO DO EVERY ONE. BE SURE TO MAKE A COMPLETE SENTENCE.

1.	I like - - - - - - - - - - - - - -	21.	I failed - - - - - - - - - - - - - -
2.	The happiest time - - - - - - - - - -	22.	Reading - - - - - - - - - - - - - -
3.	I want to know - - - - - - - - - -	23.	My mind - - - - - - - - - - - - - -
4.	Back home - - - - - - - - - - - - -	24.	The future - - - - - - - - - - - - -
5.	I regret - - - - - - - - - - - - - -	25.	I need - - - - - - - - - - - - - -
6.	At bedtime - - - - - - - - - - - -	26.	Marriage - - - - - - - - - - - - -
7.	Boys - - - - - - - - - - - - - -	27.	I am best when - - - - - - - - - -
8.	The best - - - - - - - - - - - - -	28.	Sometimes - - - - - - - - - - - -
9.	What annoys me - - - - - - - - -	29.	What pains me - - - - - - - - - -
10.	People - - - - - - - - - - - - -	30.	I hate - - - - - - - - - - - - - -
11.	A mother - - - - - - - - - - - -	31.	This school - - - - - - - - - - - -
12.	I feel - - - - - - - - - - - - - -	32.	I am very - - - - - - - - - - - -
13.	My greatest fear - - - - - - - - -	33.	The only trouble - - - - - - - - -
14.	In high school - - - - - - - - - -	34.	I wish - - - - - - - - - - - - - -
15.	I can't - - - - - - - - - - - - - -	35.	My father - - - - - - - - - - - -
16.	Sports - - - - - - - - - - - - - -	36.	I secretly - - - - - - - - - - - -
17.	When I was a child - - - - - - - - -	37.	I - - - - - - - - - - - - - - - -
18.	My nerves - - - - - - - - - - - -	38.	Dancing - - - - - - - - - - - -
19.	Other people - - - - - - - - - - -	39.	My greatest worry is - - - - - - - -
20.	I suffer - - - - - - - - - - - - -	40.	Most girls - - - - - - - - - - - -

FIG. 1. The Rotter Incomplete Sentences Blank.

selected and intensively studied, is of more use than a large group, poorly classified and studied only cursorily. In this research, subjects were obtained both from individual and group administrations, after which experienced psychologists were asked to classify each one as maladjusted, normal, or questionable.[3] The following general definition of maladjustment was used:

> Clearly maladjusted individuals are those students who are definitely in need of personal counseling and therapy because of personality problems. Normal students are defined as ones who would not usually be considered to be in need of treatment.

In addition to the general rating of adjustment, the psychologists rated each subject on a three-point scale in the following areas of adjustment: family, social, sexual, health, and vocational-educational. For this purpose rating scales based on behavioral examples were used.

Subjects rated in this matter came from two sources, those who had come to the Ohio State Psychological Clinic and those in small remedial psychology courses. Combined, they formed the criterion groups. Information about them was varied, ranging from that gained through intensive clinical interviews to that obtained from test data, observations in the psychology courses, and brief interviews. In the female group there were 42 college women, 20 classified as "definitely of normal adjustment", 15 as "clearly maladjusted", and 7 as "questionably adjusted". Of the 33 male students, 13 were rated as normal, 15 as maladjusted, and 5 of questionable adjustment.

On the basis of the criterional cases, scoring booklets for both males and females were developed containing examples for each item. Using the Rotter-Willerman method of scoring, all responses were divided into three categories: conflict, or unhealthy; positive, or

[3]Grateful acknowledgment is made to Dr. John Kinzer, Dr. Emily Stogdill, Dr. Mary Alice Price, and Mr. Robert Morton for providing cases for this and the latter parts of the study.

healthy; and neutral. Conflict responses were those associated with a maladjusted or unhealthy frame of mind, while positive answers were of a healthy, hopeful or humorous nature. A continuum with three levels of conflict responses (C3, C2, C1) and positive responses (P3, P2, P1) was established. With neutral responses (N) at the center, a seven-point scale was obtained, ranging from 0 to 6. The extreme positive score (P3) was scored as zero, the extreme conflict score (C3) as six. Scoring examples were chosen on the basis of the frequency of occurrence of certain kinds of responses in criteria cases. In addition, the general clinical experience of the authors and specific experience with the Army form of the test contributed to the evaluation of the degree of maladjustment indicated by a given response.

Replies for each category were varied. Conflict responses ran the gamut from hostility reactions, pessimism, symptom description, hopelessness and suicidal wishes, to statements of unhappy experiences, while positive answers might be humorous or flippant remarks, optimistic expressions or acceptance statements. Generally speaking, the neutral responses were on a simple descriptive level. Some replies, such as "A mother . . . has children" are to a large extent merely repetitious of the stimulus. Other common neutral answers are stereotypes, catch phrases, song titles, and essentially meaningless and fragmentary completions. Examples are: "Back home . . . on the farm"; "I regret . . . that I have but one life to give for my country"; "Boys . . . will be boys"; "People . . . are funny"; "Sometimes . . . I feel like a motherless child"; "When I was a child . . . I spake as a child"; etc. Though it might be hypothesized that such responses, as well as humorous or flippant remarks, would be highly correlated with maladjustment, it was found that avoidance answers of this type correlated at least as well with good adjustment as with bad. This finding is in agreement with the earlier Convalescent Hospital study.

In an attempt to discover and eliminate common scoring problems, once the manual had been developed, the papers of 16 additional females and 20 males, whose adjustment was not as adequately determined as the criterion group, were examined and rated. Frequent responses not included previously in the scoring manual were used to expand it to its final form. In this step the adjustment of the subjects was not as well established and the basis for placing a response in a given category rested primarily on the judgment of the authors, based on a study of the whole record. Altogether, the protocols of 58 female and 53 male subjects were used in developing the scoring manuals.

The actual scoring examples for item 35 are given below for illustrative purposes.

MALES

My father - - - - - -

C_3. promises many things and never keeps them; is the male responsible for my existence; wasn't very good; was a fool; and I have many arguments; is an alcoholic.

C_2. and I never were buddies; is in pretty bad shape; is hard to understand; is stern.

C_1. is in bad health; is good to me, but we have little spiritual communion; cannot supply me with everything; never had much of a chance; lives in - - -; is proud; is sensitive.

N. is home; is a salesman; is dead; had all his teeth pulled; is a hard worker; is living.

P_1. is good to me; is very intelligent, though not highly educated; is an idol to me; is an excellent mechanic; is a good hard-working man; is the kindest, most honest man I have ever known.

P_2. is extremely caustic and reactionary, but I love him; is OK; is all right; is a good man.

P_3. is the greatest dad in the world; is a swell guy; is a good joe; is a good companion.

FEMALES

My father - - - - - -

C_3. hasn't been home since I was 12 years old; still frightens me; is a stranger to me; is alcoholic.

C_2. isn't going to change and I wish he would; is not ambitious enough; is pretty strict; and I were never too close; has always made us work very hard at home; is (was) so good to me; antagonizes me; is angry with me; isn't sociable at home; is dead but I think of him a lot.

C_1. is really a good person but does not know how to warm up to people; seems more understanding than my mother; I wonder if I'll ever meet anyone as grand; worries about me too much; is a very quiet man; is the best man I know; is dead (with no feeling expressed about father's death.)

N. is a successful business man; is a - - - (occupation); raised a large family; is hard-working; is in - - - (place).

P_1. is quite a character; cultivated my interests in sports; (activity with father, e.g., and I discuss current events daily); is OK; is all right; has my respect and admiration; is a good man; is wonderful.

P_2. is a very handsome and intelligent metallurgist; has a complete head of hair - - hurrah!; is very nice.

P_3. has a wonderful sense of humor; is very congenial; is a lot of fun; is a good guy.

VALIDATION OF THE SCORING MANUALS

Cases used for validation of the female manual were taken from three kinds of sources: A, classes in effective study; B, classes in mental hygiene; and, C, a limited number of cases individually declared, by competent clinicians, to be adjusted or maladjusted. In Groups A and B the instructors were asked to rate every member of the class as either maladjusted or adjusted, i.e., as needing personal counseling or not. Thus forced to classify all subjects he possibly could in one or the other category, the instructor made judgments in many cases where he was relatively unsure of his rating. A number of errors of judgment would be expected in this group, which will, in this report, be referred to as "forced choice" cases.

Inadequate numbers of maladjusted females were available so that most of the clear-cut cases were used in the development of the manual. In order to strengthen the criterion for the female group a few advanced student clinicians were asked to secure tests from any college students whom they knew to be clearly maladjusted or well-adjusted. In this fashion, six cases considered maladjusted and four well-adjusted were obtained. These ten comprise Group C.

The male subjects included the following groups: A, forced choice cases from effective study classes; B, forced choice cases from mental hygiene classes; C, self-referrals to the psychological clinic for treatment; and D, cases from the Occupational Opportunities Service, referred for personal counseling by the vocational advisors.

Two females and one male, all forced choice cases, were dropped from the initial group when further experience with the subjects convinced the instructors of the classes that they had been incorrectly classified. For the final study 82 female and 124 male records were used. The distribution of male and female cases used is given in Table I.

TABLE I

SOURCES OF SUBJECTS FOR MALE AND FEMALE CROSS-VALIDATION STUDY

Group	Females			Males		
	Adjusted	Maladjusted	Total	Adjusted	Maladjusted	Total
A	24	17	41	48	13	61
B	26	5	31	9	8	17
C	4	6	10	0	15	15
D	0	0	0	0	30	30
Total	54	28	82	57	67	124

After all cases were collected, the following procedure was applied: the papers were numbered at top and bottom by one author and all identity removed, and they were then shuffled. In this manner bias was avoided, since the adjusted and maladjusted papers were interspersed, and the scorer who had no previous familiarity with the papers, did not know whether they had been classified adjusted or maladjusted.

TEST RELIABILITY

Since it is hoped that the test is sensitive to changes in adjustment, test-retest reliability is not a good measure of reliability. Internal consistency, as measured by the split-half method, is of some interest in providing an estimate of reliability, but is not strictly applicable since the items cannot be said to be measuring the same thing or to be graded for difficulty. In a test of this kind where scoring must be done from examples, and it is possible for some subjectivity to appear in scoring (although each item is generally supposed to be scored independently of all others), inter-scorer reliability is perhaps of greatest interest. In light of this, consequently, two methods were used to estimate the reliability of the test. (1) For reliability of scoring, two judges, both advanced students in clinical psychology who were thoroughly familiar with the test and methods of scoring, independently scored the same 50 male records and 50 female records and their inter-scorer reliability was determined by the product-moment correlation. (2) The reliability of the instrument was estimated by the split-half method. The test was divided into equivalent items on an *a priori* basis and the correlation between halves computed by the product-moment method for both the males and females.

The correlations, standard errors, means, and standard deviations for interscorer reliability for 50 female and 50 male records are given in Table II.

TABLE II
INTER-SCORER CORRELATIONS ON THE INCOMPLETE SENTENCES BLANK

Scorer	r	SE	M	SD
FEMALES (N = 50)				
X			126.7	19.82
	.96	.01		
Y			125.0	21.17
MALES (N = 50)				
X			128.8	15.86
	.91	.02		
Y			125.1	16.90

When one judge scored 71 female records and the test was divided into halves on an *a priori* basis, the product-moment correlation between the equivalent halves was found to be .83 as corrected by the Spearman-Brown prophecy formula.

Similarly, one judge scored 124 male subjects and determined equivalent halves on an *a priori* basis. The correlation between the halves was found to be .84 corrected by the Spearman-Brown prophecy formula.

TEST VALIDITY

The main measure of test validity was obtained by determining the biserial coefficients between scores on the Incomplete Sentences Blank and the classification of adjusted or maladjusted. A secondary clue to validity could be obtained by comparing the maladjusted cases rated by the forced choice technique with the obviously more serious cases who had found their way into clinics or sought help in some form. Since the groups involved are not large enough to make a statistical comparison for females this was done only for males.

The actual distribution of all cases is given in Tables III and IV. Group means and biserial validity coefficients

are given in Tables V and VI.

The difference between the means of the forced choice maladjusted males and the clinic cases was significant beyond the .01 level of confidence.

Although the forced choice method in one of the classrooms (Group B) showed little differentiation, the success of the test with other groups produced relatively high validity coefficients. The relatively small number of selected female cases which showed no overlap in score suggests that the lower biserial was due to the smaller percentage of clearly maladjusted girls in the maladjusted group.

By using a cutting score of 135 it is seen (Table III) that the test correctly identifies 168 per cent of the maladjusted females and 180 per cent of the adjusted females. With the same cutting score (Table IV) 169 per cent of the maladjusted males and 189 per cent of the adjusted males are correctly placed.

TABLE III

DISTRIBUTION OF SCORES ON THE INCOMPLETE SENTENCES BLANK FOR FEMALES*

Score	Adjusted (54)	Maladjusted (28)
180-184		(1) A
175-179		
170-174		
165-169		
160-164		(4) A B C C
155-159		(4) A A A C
150-154		(2) C C
145-149	(2) A B	(3) A A C
140-144	(5) A A A A B	(4) A A A A
135-139	(4) A B B B	(1) B
130-134	(5) A A A B B	(1) A
125-129	(6) A A A A B B	(1) A
120-124	(6) A A B B B B B	(1) A
115-119	(8) A A A A B B C C	(1) B
110-114	(5) A A B B C	(2) A A
105-109	(6) A A B B B B	(2) A B
100-104	(1) B	(1) B
95-99	(1) B	
90-94	(3) B B B	
85-90	(1) C	
80-84	(1) A	

*Letters refer to the source groups as described in the text.

TABLE IV

DISTRIBUTION OF SCORES ON THE INCOMPLETE SENTENCES BLANK FOR MALES*

Score	Adjusted (57)	Maladjusted (67)
180-184		(1) D
175-179		(5) C C C C D
170-174		(2) D D
165-169		(3) C D D
160-164		(6) A C C C D D
155-159		(8) A C C D D D D D
150-154	(1) A	(2) A C
145-149	(1) A	(1) A
140-144		(6) A B C D D D
135-139	(4) A A A A	(12) A A A B B C D D D D D D
130-134	(5) A A A B B	(3) B D D
125-129	(6) A A A B B B	(5) A A C D D
120-124	(10) A A A A A A A A A A	(8) A A B B B C D D
115-119	(9) A A A A A A A A B	(4) A C D D
110-114	(11) A A A A A A A A A B B B	
105-109	(5) A A A A A	(1) B
100-104	(2) A A	
95-99	(3) A A A	

*Letters refer to the source groups as described in the text.

TABLE V

RELATIONSHIP BETWEEN PSYCHOLOGIST'S RATING OF ADJUSTMENT AND SCORES ON THE INCOMPLETE SENTENCES BLANK FOR FEMALES

Group	N	M	SD	Bi-serial r	SE
Adjusted	54	120.15	15.64		
A	24	124.5			
B	26	117.96			
C	4	108.25			
Maladjusted	28	140.93	20.93		
B	17	140.53			
B	5	125.00			
C	6	155.33			
All students	82	127.24	19.98	.64	.10

TABLE VI

RELATIONSHIP BETWEEN PSYCHOLOGIST'S RATING OF ADJUSTMENT AND SCORES ON THE INCOMPLETE SENTENCES BLANK FOR MALES

Group	N	M	SD	Bi-serial r	SE
Adjusted	57	119.37	11.70		
A	48	118.88			
B	9	122.00			
Maladjusted	67	144.31	18.97		
A	13	137.38			
B	8	125.57			
C	16	154.5			
D	30	146.33			
All Students	124	132.85	20.30	.77	.06

EVALUATION OF RESULTS

In order to obtain an adequate sample of cases for cross-validation purposes it was necessary to draw on cases from a number of sources and by so doing to produce criteria of different strength. It was particularly difficult to obtain an adequate sample of maladjusted female cases and this appears to be reflected in lower biserial coefficient of validity. In light of the weakness of the criteria, the results appear to indicate relatively high efficiency in screening and suggest that increase in score is in some way directly proportionate to increase in conflict or maladjustment. It might be thought that the higher score is a reflection of readiness to accept help. However, the six specially selected maladjusted cases in the female group were not individuals who were known to have sought help.

Interscorer reliability for experienced scorers is high enough to suggest that scoring was reasonably objective.

The similarity of results of the present study with the previous study of personality screening in an AAF Convalescent Hospital [6] suggests that this method of scoring by example might be utilized for a number of screening problems. It appears that one of the greatest advantages of the incomplete sentence method is its flexibility. Experienced psychologists can devise special items and scoring examples for many specific problems. The method shows promise for a number of other screening problems, as in industry or the schools. It also has particular potentialities as an objective attitude scale and as such it could be used to reveal attitudes toward race, politics, industrial relations, institutions, and so forth.

GENERAL OBSERVATIONS

Further analysis of the results of this study have produced some observations of interest to the general problem measuring maladjustment. Many of these are suggestive of further research.

Generally speaking, it was found that there is a tendency for all individuals to twist responses. For example, it was noted that well-adjusted students twist answers, so that in reply to stimuli which presuppose a negative answer the subject responds in a positive or healthy manner. For example, in response to the stimulus, "What annoys me . . . ", a well-adjusted individual might reply, "people who squeeze the tooth paste in the middle"; or to "What pains me . . a blow in the solar plexis." In reply to the same stimuli, however, the poorly adjusted individuals might respond; "What annoys me . . . are people"; "I can't . . . think straight"; "What pains me . . . is my home life".

It was found also that poorly adjusted students twist responses in reply to stimuli which suggest a positive answer. For example, instead of replying to "I like . . . a great many things", as would a well-adjusted person, he responds, "I like . . . to be alone." Where a healthy individual would reply, "The happiest time . . . is now", the maladjusted subject says, "The happiest time . . . ends badly". In opposition to a positive response such as "The best . . . is yet to come", a conflict response is "The best . . . friends are animals like dogs and cats."

As has been previously stated it might be hypothesized that neutral as well as humorous or flippant remarks would be correlated to a high degree with maladjustment. However, in general, it seems that the maladjusted individuals take the test, as well as themselves, in all seriousness, the test gives them an opportunity to express their feelings of helplessness and desire for sympathy and help. In reading the papers one often gets the impression that the plaintive cry of the

poorly adjusted individual is that "no one understands me", and that these tests give some measure of outlet. On the other hand, the well-adjusted persons with their abundance of flippant and neutral responses seem to be saying in effect that their feelings are private property.

In analyzing the responses made by the maladjusted and adjusted individuals, an apparent correlation was found between the length of the response and the adjustment of the individual. The maladjusted individual often writes long, involved sentences as if compelled to express himself fully so as not to be misunderstood. On the other hand, the well-adjusted person frequently replies to the stimuli with short concise statements. On an inspection basis, with a cutting score of six or more statements that were more than one line in length, it was found that nine of fourteen such papers had been rated as maladjusted. For example, one poorly adjusted person wrote, "I feel . . . a deep sympathy for the needs and troubles of people and have a desire to help." An adjusted person wrote, "I feel . . . fine."

Another aspect of the completions appears to be related to length of response. This is a measure of the amount of emotional involvement as shown by the strength of language used, such devices as the exclamation mark and underlining. For example, "Most girls . . . around here are stuck on *themselves*!!", or "Most girls . . . are no damn good!", or "Boys . . . are mostly asses!", or "My nerves . . . are shot to hell", or "The only trouble . . . wish I had only ONE." All these examples illustrate a great degree of emotional involvement. Stein [8] has also noted the relationship of lengthy and strongly termed responses to maladjustment.

Although perseveration or frequent repetition of the same responses was not particularly frequent in this college group, experience with the test has indicated that these manifestations tend to occur more frequently in more disturbed groups of maladjusted individuals. A rigidity or perseveration score might be of value for some experimental studies.

SUMMARY

Manuals for scoring of an incomplete sentence test, adapted from an earlier one used by Rotter and Willerman, were developed for male and female college students in order to appraise degree of conflict or maladjustment. In developing the manuals 58 female and 53 male subjects were used.

Validity and reliability of the test were determined on new groups of 82 female and 124 male subjects. To prevent bias the source of the records was unknown to the scorers. Interscorer reliability of the test for two experienced scorers was .96 for the female manual and .91 for the male manual. Split-half reliability was .83 for the female manual and .84 for the male manual. However, split-half reliabilities are not strictly applicable to this type of test because of the nonequivalence of items.

The biserial validity coefficients obtained between ratings of adjusted or maladjusted and test score was .64 for females and .77 for males. It was considered that the lower coefficient for females was a function of the less adequate criterion used.

In general, the test appears to be promising for use with college students for a variety of screening and experimental problems when a measure of degree of conflict or maladjustment is required.

Received January 24, 1949.

REFERENCES

1. BELL, J. E. *Projective techniques.* New York: Longmans, Green, 1948.
2. HOLZBERG, J., TEICHER, A. AND TAYLOR, J. L. Contributions of clinical psychology to military neuropsychiatry in an army psychiatric hospital. *J. clin. Psychol.*, 1947, 3, 84-95.
3. HUTT, M. L. The use of projective methods of personality measurement in army medical installations. *J. clin. Psychol.*, 1945, 1, 134-140.
4. ROHDE, A. R. Explorations in personality by the sentence completion method. *J. appl. Psychol.*, 1946, 30, 169-181.
5. ROHDE, A. R. A note regarding the use of the sentence completions test in military installations since the beginning of World War II. *J. consult. Psychol.*, 1948, 12, 190-193.
6. ROTTER, J. B. AND WILLERMAN, B. The incomplete sentences test as a method of studying personality. *J. consult. Psychol.*, 1947, 11, 43-48.
7. SHOR, J. Report on a verbal projective technique. *J. clin. Psychol.*, 1946, 2, 279-282.
8. STEIN, M. I. The use of a sentence completion test for the diagnosis of personality. *J. clin. Psychol.*, 1947, 3, 46-56.
9. SYMONDS, P. M. The sentence completion test as a projective technique. *J. abnorm. soc. Psychol.*, 1947, 42, 320-329.

VALIDITY OF SHORT FORMS OF THE WECHSLER-BELLEVUE SCALE FOR MENTAL DEFECTIVES

BY MILTON COTZIN AND JAMES J. GALLAGHER
SOUTHBURY TRAINING SCHOOL, SOUTHBURY, CONNECTICUT

THERE has been a trend in recent years towards the use of abbreviated forms of the Wechsler-Bellevue scale in the measurement of intellectual capacity. The popularity of short forms can be traced to the practical situation which occurred during the last war when there was a great need for rapid individual testing in military centers where large numbers of men were screened or classified by psychiatrists and psychologists. The pressure of such work did not allow time for the administration and analysis of a full Wechsler. Since its use was restricted in such clinical situations, short forms seemed necessary for the determination of the general intellectual level.

The Wechsler-Bellevue scale [8], which (a) gives a reliable measure of adult intelligence, (b) is considered to be of significance diagnostically, and (c) is easily subdivided, has been the object of several abbreviated scales. Each of the proposed short forms has attempted more or less to incorporate the various factors which would make an initial brief test sound. Such factors include the following objectives and criteria: (1) a reduction in time for administration, (2) an accurate measurement of intelligence, (3) objectivity and ease of scoring, and (4) the retention of items of diagnostic significance. The disadvantages of the recent movement towards the use of shorter and shorter tests for clinical practice have been aptly stated by Patterson [5]. They involve the expectation that a judgment of intellectual functioning and emotional stability can be reached in a few minutes, the discrediting of psychology because of unreliable or invalid results, the decrement in the reliability of the sample of the subject's behavior because of the short time for testing, and the reduction in the aspects of functioning which can be observed and measured. However, the value of abbreviated forms is still high in the overcrowded and understaffed institutions and clinics throughout the country, in the testing programs of schools, in the hasty examinations in the courts and industry, and elsewhere, where large numbers of people are referred routinely for psychological reasons. Moreover, in such situations the time saved by the use of a short scale can be utilized by the clinician in exploring personality dynamics or special abilities which may be of greater importance than the general intellectual level.

The proposed short forms have been standardized and validated on special groups such as male prisoners [2, 3], male naval recruits [1, 4, 7], student nurses [6], psychotic patients [6], and other severely maladjusted individuals [5]. The Wechsler-Bellevue IQ's in these groups ranged from a low of 39 to a high of 147. A brief description of the previous populations tested, along with the proposed abbreviated scales,[1] is in order.

Gurvitz utilized in his study (D-PA) [3] a total of 523 male adult prisoners with an age

[1]The ten Wechsler subtests (vocabulary excluded) will be referred to throughout by letters with all verbal tests indicated by single letters and all performance tests by double letters.

range of 17 to 64 and an IQ range 42 to 147. These prisoners were referred for either psychometric or psychiatric reasons and 250 of them were maladjusted mentally and physically. Cummings' group (C-A) [1] was made up of 420 white male recruits of questionable intellectual quality. The verbal IQ range here was 48 to 98 with a mean of 73.34 and a standard deviation of 11.80. The age range was 17 to 40 with a mean of 21.26 and a standard deviation of 5.06. The group did not include any cases of psychosis, organic brain disease or epilepsy. In Rabin's studies (C-A-S) [6] he utilized two groups: (a) 92 student nurses with an age range of 19 to 25 and an IQ range of 85 to 130, (b) 200 state hospital patients with an age range of 15 to 36 and an IQ range of 39 to 122, these patients being mainly psychotic although a few were neurotic. There were 128 males and 72 females.

Springer (C-A-S) [7] used 100 naval recruits aged 20 to 24 with an IQ range of 67 to 119 with a mean of 82.95 and a standard deviation of 12.35. These men had been referred for an examination because they were suspected of being mentally retarded. Geil (C-S-D-BD and also I-PC-PA-DS) [2] employed 250 unselected cases at the Medical Center for Federal Prisoners, Springfield, Missouri. This population consisted mainly of psychotic, physically handicapped, and otherwise maladjusted individuals whose IQ range was from 40 to 132 with a mean of 90.30 and a standard deviation of 20.66. No age ranges are given. Patterson in his study (a validation of the above short forms and a proposal of scales to include vocabulary), [5] used 50 patients in an overseas army hospital specializing in the care of psychiatric patients. The mean age of the group was 23.4 with a standard deviation of 4.6 and a range of 19 to 37. Full scale IQ's ranged from 51 to 116 with a mean of 77 and a standard deviation of 18.76. This group included major neuroses and psychoses and several organic brain diseases or trauma.

PURPOSE

The purpose of this study is to test the validity of these various short forms for institutionalized mentally defective individuals. Although previous writers have included in their test populations individuals whose measured intellectual level falls in a so-called defective range, it is considered (1) that a low IQ does not necessarily mean a mental defective, (2) that the validity of a short form standardized on a population ranging from below IQ 70 to above IQ 100 does not hold necessarily for a homogeneous group of mental defectives, and (3) that short forms standardized upon special groups as psychotics, naval recruits, nurses, or prisoners may not necessarily be valid for well-adjusted individuals with retarded intellectual capacities.

Wechsler [8] has pointed out that the test pattern for mental defectives is different clinically from the pattern for other clinical groups. Furthermore, it is generally known that intellectual levels above borderline show small differences between the verbal abilities and the performance abilities of the subject. However, with mental defectives, for the most part, there is a great discrepancy between the verbal IQ and the performance IQ. A short form to be of significance for mental defectives must give, in addition to an accurate estimate of a full scale IQ, a valid representation of the defective's verbal and performance capacities. Such an abbreviated Wechsler scale not only will provide more adequate clinical information about the defective but also is of considerable value administratively. Since group tests are not feasible, a short individual scale can have considerable merit in institutions for the mentally defective where there may be pressure for routine retesting or where time is needed for other clinical studies on the newly-admitted defective who has already had his intellectual level quite well established. This study checks the previous writers' suggestions concerning the utility of their scales for the mental defective.

SUBJECTS

The subjects used in this study were 154 mental defectives in residence at

the Southbury Training School. The department of psychological service had classified each as a "high grade" or "borderline" defective. Within the institutional population, a "high grade" defective is defined as that individual who receives a full scale Wechsler-Bellevue IQ between 49 and 67 and who has failed to make a satisfactory social adjustment prior to admission; a "borderline" defective, although he obtains a full scale Wechsler-Bellevue IQ between 68 and 79, has demonstrated, nevertheless, an inability to meet the demands of society without strict supervision and guidance. "Middle" and "low" grades of deficiency (IQ's under 49) were excluded since the Wechsler-Bellevue test is considered unsuitable for these more severe levels of deficiency. Each subject had been legally judged a mental defective in accord with psychological, social, and developmental criteria, and also in respect to actual functioning. Their adjustment to institutional living, however, had been rated as "reasonably good" or "good". Epileptics, psychopaths, neurotics, psychotics, cerebral palsies, and etiologies other than familial or undifferentiated were excluded from the study group.

Table I gives the breakdown in age, verbal IQ, performance IQ, and full scale IQ for sex, etiology and classification of the group studied. The mean age of the group was 23.50 years with a range from 14 to 39 years. Defectives over 40 years of age were excluded from the group since experience at Southbury has shown that full scale Wechsler IQ's obtained on such individuals are significantly higher than the level established by many other previous measures. The full scale Wechsler-Bellevue IQ's ranged from 49 to 79 with a mean of 64.40; the mean verbal IQ and the mean performance IQ for the entire group were 64.19 and 70.98, respectively. The cases are proportionate and representative in all respects—sex, age, intellectual levels, and etiology — of the trainable population in residence at the Southbury Training School [9] and are probably representative also of similar populations in other institutions for the mentally defective.

METHODS AND PROCEDURES

The ten subtests of the Wechsler-Bellevue scale were administered under the usual standardized conditions to 154 resident of the Southbury Training School as part of the diagnostic classifi-

TABLE I

AGE, VERBAL IQ, PERFORMANCE IQ, AND FULL SCALE IQ FOR SEX, ETIOLOGY, CLASSIFICATION, AND TOTAL GROUP OF 154 DEFECTIVES STUDIED

	Age			Verbal IQ		Perf. IQ		Full IQ	
	No.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Sex									
Male	68	21.07	5.97	63.28	9.50	70.93	11.75	63.62	7.88
Female	86	25.42	6.95	64.92	6.28	71.03	10.37	65.01	7.47
Etiology									
Familial	114	23.41	6.75	64.79	7.75	71.84	10.91	65.19	7.39
Undifferentiated	40	23.76	7.23	62.50	8.23	68.52	11.08	62.12	8.16
Classification									
High Grade	80	23.22	6.31	59.46	6.19	64.44	9.46	58.18	5.01
Borderline	74	23.80	7.45	69.31	6.33	78.05	7.78	71.12	2.90
Total	154	23.50	6.88	64.19	7.98	70.98	11.04	64.40	7.66

cation or reclassification procedures. The mean IQ and standard deviation of the verbal subtests, the performance subtests, and the full scale tests were computed as shown in Table I for the entire group and for various breakdowns as sex, classification, etiology, and age. The mean weighted score and standard deviation for each of the ten subtests were computed.

In order to make a valid assessment of each form proposed by previous authors, the procedures laid down in their original articles for obtaining full weighted scores were followed closely. In all but the Cummings report, this procedure consisted of obtaining the mean weighted subtest scores of the short form and multiplying by 10, thereby getting a predicted short scale weighted score comparable to the actual full scale weighted score. For example, in a two-item short form the sum of the weighted scores of the two items is multiplied by 5; if a four-item short scale had been used, the sum of the four individual weighted scores is multiplied by $2\frac{1}{2}$. Cummings, in his analysis of a short form for predicting a verbal IQ, first found the coefficient of correlation between the sum of his two-item weighted scores and the full scale weighted scores and then applied a regression equation to this coefficient of correlation; new equivalent weighted scores were then obtained on the basis of this regression equation. In the present validation for mental defectives of the short scales, the above methods of obtaining equivalent full scale weighted scores were adhered to.

Since a systematic analysis of the relationship between two psychometric instruments requires more than a mere determination of the degree of correlation between them, analysis was first made of differences in terms of the critical ratios between the distributions of the predicted and actual values for each short form. In any abbreviated scale where there was a high critical ratio it could be considered that the predicted score had little practical value. On the other hand, a nonsignificant critical ratio along with a high degree of correlation between the distributions might indicate predicted scores of great value. Pearsonian coefficients of correlation were run between the predicted full scale weighted scores and the actual full scale weighted scores. In addition, the new predicted weighted score for each subject on each abbreviated scale was converted into an equivalent IQ by means of the Wechsler tables taking age into account, and correlation coefficients were obtained between the predicted IQ based on the short form and the actual IQ found on the administration of the full Wechsler.

In line with the need for a more systematic analysis and in view of the practical considerations for the use of a short form with mental defectives, an analysis was made of the amount of deviation of each individual's predicted IQ score from his actual score. The mean deviation was found for each of the previously proposed short scales, and in addition, the percentages of cases in which the individual deviations ranged between zero and four, between five and nine, and also ten or more points above or below the actual IQ were computed. In the practical situation where a test-retest analysis of results is made, it is felt with mental defectives that a deviation from prior findings of four points can be considered to be of no significance as regards a change in the intellectual level or development of the subject. A deviation of five to nine points from the previously obtained IQ is also considered to be within limits of a test-retest situation, but such a deviation might be probed further in order to

determine whether an actual change clinically has taken place in the subject. In any case where there is a deviation of more than ten points from previously established results, a study of the situation should be made.

Similar procedures were followed whenever validations of the short scales were indicated for verbal or performance levels separately.

RESULTS AND DISCUSSION

The mean verbal IQ, performance IQ, and full scale IQ with their respective standard deviations have already been given in Table I. The mean weighted score and standard deviation for each of the ten subtests of the Wechsler are presented in Table II.

Here it is noted that this representative group of 154 mental defectives did best on object assembly and poorest on arithmetic. There is only one verbal item (comprehension) which is better than the poorest of the performance subtests. In general, the group did better in the performance items (mean wt. sc.=5.59) than it did on the verbal items (mean wt. sc.=3.56); this is also seen in the difference between the mean performance IQ (70.98) and the mean verbal IQ (64.19). However, in treating the verbal items separately from the performance items, following Wechsler's analysis of pattern [8], one sees that the verbal subtests do not deviate from each other except for arithmetic, and the performance items do not deviate very much from each other except for the high object assembly results.

TABLE II

WEIGHTED SCORE MEANS AND STANDARD DEVIATIONS OF THE TEN WECHSLER SUBTESTS

(Full Weighted Score Mean = 4.57; N = 154)

Subtest	Mean	SD
Information	3.78	1.75
Comprehension	5.02	1.95
Digit Span	3.76	2.42
Arithmetic	1.23	2.02
Similarities	4.02	2.07
Picture Completion	5.50	2.47
Picture Arrangement	4.79	2.40
Block Design	5.01	2.54
Object Assembly	7.12	3.54
Digit Symbol	5.52	1.98

TABLE III

MEANS AND STANDARD DEVIATIONS OF FULL SCALE WEIGHTED SCORES PREDICTED FROM SHORT FORMS AND CRITICAL RATIOS BETWEEN THE PREDICTED AND FULL SCALE WEIGHTED SCORES

	Full Scale Wt. Sc.	C-A Cummings	C-A-S Rabin	D-PA Gurvitz	C-S-D-BD Geil	I-PC-PA-DS Geil
Mean	45.76	31.17	34.42	43.18	44.42	48.97
SD	11.21	13.88	13.55	17.78	13.99	12.72
CR		10.20	8.04	1.42	.92	2.36

TABLE IV

MEANS AND STANDARD DEVIATIONS OF FULL SCALE IQ'S PREDICTED FROM SHORT FORMS AND CRITICAL RATIOS BETWEEN THE PREDICTED AND FULL SCALE IQ'S

	Full Scale IQ	C-A Cummings	C-A-S Rabin	D-PA Gurvitz	C-S-D-BD Geil	I-PC-PA-DS Geil
Mean	64.40	53.94	56.14	62.36	63.57	66.70
SD	7.66	10.72	9.79	13.02	9.95	8.64
CR		9.87	8.26	1.67	.80	2.47

Tables III and IV compare the average results found with mental defectives using the previously proposed short forms and give the critical ratios between the distribution of the predicted scores and the true full scale scores. There is little difference, of course, between results employing weighted scores or IQ's as data.

Table V shows the coefficients of correlation for the full scale weighted scores and the abbreviated weighted scores as found with mental defectives. This table also gives the correlations obtained by the original authors on their populations and those found by Patterson who employed the same abbreviated scales in his sample. Correlations between predicted IQ's and actual IQ's were also obtained but differ little from the correlations shown in Table V for the weighted scores. The latter are presented since previous authors utilized such correlations and comparisons are therefore more valid.

TABLE V

PEARSONIAN COEFFICIENTS OF CORRELATION BETWEEN SHORT FORM WEIGHTED SCORES AND FULL WEIGHTED SCORES

Short Form	N	r with Full Wt. Sc.
C-A		
Cummings	420	.....
Patterson	50	.851
Southbury	154	.532 ± .038
C-A-S		
Rabin	92	.80
Rabin	200	.956
Springer	100	.92
Patterson	50	.890
Southbury	154	.588 ± .035
D-PA		
Gurvitz	523	.90
Patterson	50	.815
Southbury	154	.746 ± .024
C-S-D-BD		
Geil	250	.966
Patterson	50	.936
Southbury	154	.835 ± .016
I-PC-PA-DS		
Geil	250	.952
Patterson	50	.948
Southbury	154	.741 ± .024

In general, it is noted that the correlations obtained with a sample of defectives are below those found in other samples. This is due in part to the restriction in the range of the scores in the present population. However, comparisons among the abbreviated scales and between the defective population and the other populations are permissible.

A final criterion for the validity of short scales is the analysis of the amount of deviation between the predicted IQ's and the actual IQ's. Such an analysis was made in terms of the average deviation for the entire group studied and the percentage of cases that differed between 0 and 4 IQ points, between 5 and 9 points and more than 10 IQ points above or below the original full scale IQ. Table VI gives these values for the abbreviated scales studied.

TABLE VI

MEAN IQ DEVIATION AND PERCENTAGE OF CASES DEVIATING 0-4, 5-9, AND MORE THAN 10 IQ POINTS ABOVE OR BELOW THE FULL SCALE IQ

Short Form	Mean Deviation (IQ points)	Percentage Deviating 0-4	5-9	More Than 10 IQ
C-A	12.05	16.9%	24.0%	59.1%
C-A-S	9.58	26.0	27.9	46.1
D-PA	7.49	34.4	35.1	30.5
C-S-D-BD	4.43	59.7	28.6	11.7
I-PC-PA-DS	4.87	56.5	33.8	9.7

The two-item scale advocated by Cummings as a "special screening device for assessing the type of intellectual ability demanded by service conditions" is not applicable for prediction of full scale IQ's of mental defectives. The predicted mean weighted score and IQ fall much below the actual means and there are clearly significant critical ratios between the distributions. These are to be

expected since this scale is made up of verbal items only and defectives are generally lower in verbal than in performance ability. In addition, the correlation between this form and the full scale is small; the mean deviation is 12 points; and 59 per cent of the cases deviate more than 10 points above or below the actual IQ with 34 per cent deviating by more than 15 points. It must be pointed out that this abbreviated form had originally been employed for a predicted verbal IQ; it has, however, been used by others for comparison on a full scale basis.

In view of the original intent, Cummings' abbreviated scale was tested out on the verbal abilities of the present population, and now a more favorable picture is seen. The correlation between the C—A scale and the Wechsler verbal scores is .831. A new regression equation, $Y = 1.94\ X + 5.70$, was computed for the verbal data and equivalent verbal IQ's found for the predicted weighted scores. Here the mean deviation becomes 4.17 IQ points and 67.5 per cent of the cases deviate 0—4 points, 26.0 per cent deviate 5—9 points, and only 6.5 per cent deviate more than 10 IQ points. Whereas the C—A scale falls short in its prediction of full scale levels and should therefore not be used as a Wechsler full scale screening or measuring device, it is quite good for verbal determinations.

The Rabin short form (C-A-S), which proved itself quite successful on other clinical groups, is found to be inapplicable to defectives. Here again the predicted means fall much below the actual, and the critical ratios show significantly different distributions. Like the Cummings scale, this abbreviated form contains only verbal subtests and thus cannot represent adequately the intellectual levels of mental defectives. Furthermore, one of its items is the arithmetic subtest which is handled extremely poorly by a defective population; as noted in Table II the mean weighted score for arithmetic is 1.23 or 2.5 weighted units below the mean of any other subtest. In addition, there is a small correlation between scores on this short scale and the full scale; the mean deviation of the individual cases is 9.5 points; and this scale predicts almost half of the cases beyond the 10 point limit of a test-retest criterion, with approximately 20 per cent of the cases deviating more than 15 IQ points from the full scale values. Springer [7] used the Rabin form on a group which had a full scale IQ range of 67 to 119 and a mean of 82.95; the present study employed a full scale IQ range of 49 to 79 with a mean of 64.40. The poor results found with this abbreviated scale point out the inherent danger of standardizing a scale successfully on one sample of the population and then attempting to use it upon a different clinical sample.

The D—PA scale of Gurvitz seems on the surface to hold up well for defectives in terms of the approximation of the means and also the predicted distribution of scores which is not significantly different from the original scores. However, further analysis indicates that this scale should not be considered an accurate measuring device for defectives. Its correlation is .746; its mean deviation is 7.5 IQ points; and, furthermore, it fails in its predictive efficiency in view of the large number of cases (30 per cent) which deviates by more than 10 IQ points from the original. Gurvitz claims that his scale provides good discrimination in those categories where it is most needed, that is, in the IQ's between 40 and 70; present results with institutionalized mental defectives do not substantiate this claim. Although this short form contains both a verbal and a performance subtest, it

has rather dubious diagnostic value because of its brevity of clinical observation.

The Geil scale consisting of one verbal and three performance subtests has a distribution of scores which differs significantly from the full scale scores. It shows about the same amount of relationship as did the two-item Gurvitz form. On the other hand, this four-item scale predicts individual IQ's better than the other proposed forms; less than 10 per cent of the cases deviate by more than 10 points. However, the mean deviation is 4.9 points and most of the individual discrepancies are overpredictions.

Utilizing all of the criteria for an evaluation of the validity with defectives of previously proposed short forms of the Wechsler, it is found that the Geil scale containing comprehension, similarities, digit span, and picture arrangement is the most applicable for full scale prediction. This scale has a mean closest to the original and a nonsignificant critical ratio. It correlates .84 with the full scale weighted scores, and the mean deviation of its predicted IQ's is 4.4 points, with only 12 per cent of the individual cases deviating by more than 10 IQ points from the full scale values.

As has been pointed out in a preceding section, it is considered that a short form of the Wechsler to be significant for mental defectives must not only be highly valid for full scale IQ prediction, but must also give an accurate estimation of the defectives' verbal and performance capacities independently. Therefore, correlations were run between the Geil scale (C-S-D-BD) and the defectives' verbal weighted scores and between it and the performance weighted scores. The Pearsonian *r's* were found to be .782 and .512, respectively. Although this scale has more value than other proposed forms for prediction of full scale IQ's for defectives, it is not closely related to separate verbal or performance determinations.

SUMMARY

Previously proposed Wechsler-Bellevue abbreviated forms, which had been standardized on various clinical groups, have suggested their applicability to measurement of intellectual levels in the defective range. This paper analyzes the validity of such scales for a representative group of 154 institutionalized mental defectives.

Analysis of data in terms of critical ratios, correlations, mean deviations and individual discrepancies between short form scores and actual scores reveals the following:

1. Cummings' scale is not valid for full scale prediction but is valuable as a device for assessing verbal abilities; however, a Wechsler verbal IQ only is insufficient for a diagnosis of mental deficiency.

2. Gurvitz' scale is not considered an accurate measuring device for defectives and has rather dubious diagnostic value because of its brevity of clinical observation.

3. Rabin's short form which proved itself quite successful on other clinical groups is found to be inapplicable to defectives.

4. One of Geil's abbreviated scales has apparent value for prediction of full scale IQ's for defectives. However, any measure of a defective's capacities must include not only a full scale determination but also an assessment of verbal and performance abilities, since there is frequently a great discrepancy between these two aspects of functioning. Geil's scale does not show a close relationship to verbal or performance abilities, when these are treated separately.

5. The results found in this study's evaluation of the above and other writ-

ers' validations of short forms point out the inherent danger of standardizing a scale successfully on one sample of the population and then attempting to use it on a different clinical group.

6. There is a need for a Wechsler-Bellevue short scale which will meet not only the criteria for a sound brief test but which will also give valid predictions of verbal, performance, and full scale IQ's of mental defectives.

Received January 11, 1949.

REFERENCES

1. CUMMINGS, S. B., JR., MACPHEE, H. M., AND WRIGHT, H. F. A rapid method of estimating the I.Q.'s of subnormal white adults. *J. Psychol.*, 1946, **21**, 81-89.
2. GEIL, G. A. A clinically useful abbreviated Wechsler-Bellevue test. *J. Psychol.*, 1945, **20**, 101-108.
3. GURVITZ, M. S. An alternate short form of the Wechsler-Bellevue test. *Amer. J. Orthopsychiat.*, 1945, **15**, 727-733.
4. HUNT, W. A., KLEBANOFF, S. G., MENSH, I. N., AND WILLIAMS, M. The validity of some abbreviated individual intelligence scales. *J. consult. Psychol.*, 1948, **12**, 48-52.
5. PATTERSON, C. H. A comparison of various "short forms" of the Wechsler-Bellevue scale. *J. consult. Psychol.*, 1946, **10**, 260-267.
6. RABIN, A. I. A short form of the Wechsler-Bellevue test. *J. appl. Psychol.*, 1943, **27**, 320-324
7. SPRINGER, N. N. A short form of the Wechler-Bellevue intelligence test as appiled to naval personnel. *Amer. J. Orthopsychiat.*, 1946, **16**, 341-49.
8. WECHSLER, D. Measurement of adult intelligence. (3rd Ed.) Baltimore: Williams & Wilkins, 1944.
9. YANNET, H. Diagnostic classification of patients with mental deficiency. *Amer. J. Dis. Child.*, 1945, **70**, 83-88.

A SUGGESTED METHOD FOR ANALYZING AND EXTENDING THE USE OF BELLEVUE-WECHSLER VOCABULARY RESPONSES

By REVA A. GERSTEIN
UNIVERSITY OF TORONTO

FOR many years, the literature in the field of intelligence testing has been laden with psychometric pattern research. Underlying this approach is the hope that differential patterns of response to intelligence tests will be found, which will aid in the diagnosis of functional and organic mental disturbances.

The following paper suggests an approach which places its emphasis upon the relationship between the use of language (as manifested in a vocabulary response) and the underlying cognitive process. Such research, in lieu of psychometric patterns, orients the field of intelligence testing to a frame of reference concerned with seeking differential levels of intellectual functioning as manifested in the development of children's vocabulary definitions, in normal adults and adults suffering from mental disturbances.

Wechsler [8] has pointed out that, "in defining a word, a subject gives more than its mere meaning." In many instances, he tells us a good deal about himself, or at least about the quality and character of his thought processes. These facts may not be immediately apparent, but can be frequently brought to light by analyzing the formal and contextual aspects of the subject's response. There is an obvious difference in the reasoning ability between two adults, one of whom defines "donkey" as "an animal," and the other who defines it in such terms as "it has four legs." Wechsler does not, however, suggest a systematic technique for vocabulary appraisal.

Vocabulary responses have been considered as one of the most stable subtests in the Bellevue-Wechsler Intelligence test. Both Wechsler [8] and Rapaport [5] have emphasized its resistant deteriorative characteristics, and have used this fact as a base line for deteriorative indices.

However, some authors have questioned the rigid stability of vocabulary responses. Yacorzynski [9] points out, in his criticism of the Babcock test [4] which was based upon the vocabulary in the old Stanford-Binet form, that "it becomes clear now why the vocabulary does not deteriorate in proportion to the other behavior measured by the Babcock test. The deterioration makes use of only the end results in the definition of words, and therefore, superficially, it appears as if the vocabulary remains unaffected by general deterioration." The suggestion is made that perhaps the mode of definition is changed, even though a passing score for each definition is still possible.

Goldstein (2) has been very much interested in the method or *modus operandi* used to attain a specific result. In his work, he found that certain brain lesions do not destroy the ability to solve a problem, but do affect the method by which the problem may be solved.

Seashore [7] has pointed out that work methods are too often neglected as

an underlying factor in individual differences.

Krechevsky [4] has presented pertinent experimental material in his work with rats. He showed that rats with cortical lesions who were presented with two possible solutions of a maze problem were unable to use a difficult solution even though it was more efficient but clung to a less efficient, lengthier solution.

For research persons interested in the effects of brain disturbances (organic or functional) on the *modus operandi*

TABLE I

ANALYSIS OF WECHSLER VOCABULARY (NOUNS ONLY) RESPONSES

STIMULUS	DESCRIPTIVE (D)	FUNCTIONAL (F)	CATEGORICAL (C)
1. Apple	red and sweet	something to eat	fruit
2. Donkey	has long ears	ride on him	animal
4. Diamond	a ring	it's valuable	gem ; precious stone
5. Nuisance	a bad boy	some one who bothers you	pest
6. Fur	hairy stuff	they make coats from them	skin, pelt
7. Cushion	something soft and comfortable	you sit or lie on it	pillow
8. Shilling	a penny	something to spend	foreign coin ; British
10. Bacon	narrow strips with eggs	something you eat	meat from pig
11. Nail	hard skin ; thin metal	something to hammer	spike
12. Cedar	smelly stuff	for chests or boxes	evergreen ; kind of wood
13. Tint	a dob of paint	to paint or brighten	a dye ; a (blended) color
14. Armory	soldier's club	where they train soldiers	military building for training soldiers and storing arms
15. Fable	story about animals	a story where animals talk	tale with moral
16. Brim	outside ; top	to get to the outside or top of something	rim ; edge
17. Guillotine	a knife	used to chop heads off	form of execution
18. Plural	a number ; two	to double anything	more than one
20. Nitroglycerine	yellow, noisy	used in ammunition	an explosive
21. Stanza	a line	to divide a poem	verse or group of lines in a poem
22. Microscope	glass, black machine	something that enlarges	magnifying instrument
23. Vesper	twilight, night	to ring the church bells at evening	evening prayer
24. Belfry	tower,bell, church	the ringing of a bell	church bell-tower
26. Affliction	a disease, headaches	something which inconveniences you	misfortune, grief
27. Pewter	dark silver	make dishes out of it	alloy ; metal used for dishes
28. Ballast	stones	to balance ships	weight used to steady ships
29. Catacomb	cave ; hole	Egyptians used it	underground cell used by Christians
30. Spangle	something dotted, shiny	used to make a dress fancier	glittering substance used for decoration
31. Espionage	Mata Hari	like a spy	act of spying, spy system
33. Mantis	fly	hops in grass	insect, grasshopper-like
34. Hara Kiri	knife in one's stomach	to kill oneself Japanese style	disembowelling, Japanese suicide form
35. Chattel	slave	to own or posses	legal term for personal property
37. Amanuensis		someone who writes	a scribe or secretary
38. Proselyte	a disciple	a man who changes his religion	convert
39. Moiety		some amount or part	a half ; small portion

used in defining a word, detailed vocabulary analysis should present a stimulating challenge.

RATIONALE

In a study by Reichard, Schneider, and Rapaport [6], the Goldstein-Weigl [3] Color Sorting Test was used as an instrument to study concept formation in children. The authors concluded that three methods of forming concepts were indicated: (a) the concretistic, (b) the functional, and (c) the conceptual.

The present study will attempt to transplant this hypothesis. Thus we will assume that the same three methods used to solve a performance problem may be used in classifying definitional levels evoked by a verbal stimulus.

HYPOTHESIS ONE

Verbalizing our knowledge of our experiential world may be indicated by the way in which we define nouns. Definitions may be given by three different methods.

(a) *The concretistic or descriptive method.* This represents the most primitive level of verbalized definition. It implies memory of an object at the concrete, sensory level. For example, "an apple is red." The "red" refers to a visual memory response to a concrete apple. "A donkey is a thing with four legs," represents a similar descriptive definition. Table I, column 1, entitled "Descriptive, D," presents examples of descriptive definitions for most of the nouns appearing in the Bellevue-Wechsler Vocabulary subtest.

(b) *The functional or usage method.* This represents a more complex method of verbalizing a definition. A word is defined by the subject recalling the use to which that object was put in the past, e.g., "an apple is something you eat," or "a donkey is something you ride on." Table I, column 2, headed "Functional (F)" presents examples of functional definitions for all nouns in the Wechsler Vocabulary subtest.

(c) *The categorical or conceptual method.* This represents the abstract method of definition. A concept involves an "abstract" attitude [9] on the part of the subject. Goldstein's research on performance tests has indicated the inability of persons suffering from certain brain lesions to assume this attitude. "An apple is a fruit" is a conceptual definition whose *modus operandi* is more complex than the descriptive or functional methods.

The Reichard, Schneider and Rapaport study using performance tests [6] concluded that children think at a descriptive or concrete level until the age of 8. Between 8-9 years the functional level reaches its peak, and at 11 years of age the conceptual classification seems well established. We can only hazard a guess at the moment, as to whether the indications of the cognitive process as demonstrated by a performance test would be similar to the results found using a vocabulary definition analysis. It points out the need for a simultaneous research project studying the response of a large number of children at every age level to ascertain the average age at which children shift their level of definition.

Goldstein [3] has indicated a precaution which must be taken by tester's too readily placing a definition at the categorical level. His excellent example of a child calling a *rose* a "flower" demonstrates that "flower" to this child was not a concept but rather a concrete name-word, such as "rose," or "pansy." Psychologists must be on guard to distinguish between "an apple is a fruit," where "fruit" is used as a true concept, as different from a mental defective's definition "an apple is a fruit," where the subject has never learned that there are apples, pears, bananas, etc. This difficulty may be overcome when a large enough sample of words are used in the analysis.

Table I, column 3, under the heading "Categorical (C)" presents examples of conceptual or categorical definitions for the Bellevue-Wechsler vocabulary subtest.

Method

(A) *Testing for Definitional Level.* Bellevue-Wechsler vocabulary responses must be recorded verbatim.

The vocabulary responses should then be compared for definitional level with those presented in Table I. If a definition best matches the descriptive level, a tick should be placed in a column D. If it matches the functional definitions, in a column F, and if it equals the conceptual or categorical level, in a column C.

If a word is defined at more than one level, the highest level of definition is ticked. For example, "an apple is a fruit which is red" is ticked under "C".

The vocabulary analysis should be made for each word which is defined.

HYPOTHESIS TWO

Comprehension level exceeds verbalization level in defining words.

(B) *Testing the Limits of Meaning by Sentence Creation.* If a subject fails to define a word in the series, but this word is followed by more difficult words which are defined, the following procedure is suggested (after the vocabulary test has been completed).

Say: "Can you make up a sentence with the word . . . . . in it?"

(C) *Testing the Limits of Meaning by Sentence Completion.* Using the same words as those selected for part B, if the subject fails to create a meaningful sentence, the following sentence completion method is suggested for deeper probing of the subject's reasoning process.

Table II presents incomplete sentences for each Bellevue-Wechsler noun.

Indicate one sentence at a time, and read aloud as the subject reads with you.

Say: "I am going to read a sentence to you. One word is missing in each. After each sentence, I will suggest three words. You tell me which word will fit the blank best, and thus give meaning to the sentence. For example: Brother is a boy; sister is a . . . . . horse, girl or tree."

TABLE II

VOCABULARY COMPREHENSION TEST

Example: Brother is a boy; and a sister is a	horse, girl, tree.			
1. An orange is round and juicy, and an . . . . is round and red.	a) apple	b) grape	c) pear
2. A . . . . . is smaller than a horse, and it has long ears.	a) cow	b) donkey	c) elephant
4. When girls become engaged, they often wear a . . . . . ring.	a) car	b) tin	c) diamond
5. When trying to sleep, a mosquito can be a	a) moth	b) bite	c) nuisance
6. In cold countries, many people wear . . . . . coats.	a) cotton	b) fur	c) paper
7. In bed, I like to rest my head on a	a) tree	b) cushion	c) box
8. In England, about 24¢ of our money, equals a	a) quarter	b) shilling	c) dime
10. For breakfast, many people eat . . . . . and eggs.	a) bacon	b) nails	c) straw
11. A carpenter hits . . . . . with his hammer.	a) glass	b) nails	c) steel
12. . . . . trees stay green throughout the entire year.	a) cedar	b) maple	c) oak
13. Many artists use . . . . . of various colors.	a) reds	b) grass	c) tints
14. During wartime soldiers are often trained in an	a) armory	b) cave	c) bridge
15. Stories with a moral are called	a) lies	b) jokes	c) fables
16. The edge of a hat is called the	c) brim	b) flag	c) nose
17. During the French Revolution, persons were beheaded by	a) electric chair	b) guillotine	c) gunshot
18. More than one is	a) plural	b) one	c) alone
20. . . . . is used in most explosives.	a) water	b)ink	c) nitroglycerine
21. A long poem is usually divided into	a) stanzas	b) sonnets	c) acts
22. To make small objects look large a . . . . . is used.	a) telescope	b) microscope	c) stethescope
23. At twilight the . . . . . prayers are said.	a) vesper	b) matinee	c) noon
24. In the . . . . . of the church, the bell is hung	a) vestry	b) apse	c) belfry
26. People who suffer from a disease have a	a) confliction	b) affliction	c) assumption
27. Dishes made of an alloy of tin and lead are	a) copper	b) pewter	c) china
28. Stones are often put on one side of a ship for	a) ornaments	b) sale	c) ballast
29. In ancient times, the Christians hid in the	a) Louvre	b) catacombs	c) Vatican
30. Many women ornament their dresses with	a) hair-pins	b) ropes	c) spangles
31. . . . . is carried on by most countries in wartime.	a) espionage	b) salvage	c) penury
33. The . . . . . is an insect which assumes a praying pose.	a) bee	b) mantis	c) moth
34. . . . . is a form of Japanese suicide.	a) suki-yaki	b) hari-kiri	c) baja-kaka
35. The legal term for personal property is	a) res	b) sine	c) chattel
37. A man's secretary acts as his	a) nemesis	b) amanuensis	c) apprentice
38. Christianity attempts to gain	a) prosecutors	b) possitors	c) proselytes
39. The . . . . . of a clan may have a great deal of power.	a) moiety	b) maroity	c) malady

SUMMARY

A systematic method for appraising and extending the limits of Bellevue-Wechsler vocabulary responses has been, herewith, presented. It is hoped that the rationale provided will stimulate further research in the field of intellectual functioning as it occurs in the normal growth of children, and in the distorted development of persons suffering from organic or functional disturbances.

Received February 2, 1949.

REFERENCES

1. BABCOCK, HARRIET. An experiment in the measurement of mental deterioration. *Arch. Psychol.*, N. Y. 1930, No. 17.
2. GOLDSTEIN, K. *The organism.* New York: American Book Co., 1939.
3. GOLDSTEIN, K. AND SCHEERER, M. Abstract and concrete behavior: an experimental study with special tests. Psychol. Monogr., 1941, No. 2.
4. KRECHEVSKY, L. Brain mechanisms concerned with problem solving. *Psychol. Rev.*, 1940, 47, 43-58.
5. RAPAPORT, D. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945. Vols. I and II.
6. REICHARD, S., SCHNEIDER, M., AND RAPAPORT, D. The development of concept formation in children. *Amer. J. Orthopsychiat*, 1944, 14, 156-161.
7. SEASHORE, R. H. Work methods: an often neglected factor underlying individual differences. *Psychol. Rev.*, 1939, 46, 123-141.
8. WECHSLER, D. *The measurement of adult intelligence.* (3rd. Ed.) Baltimore: Williams & Wilkins Co., 1947.
9. YACORZYNSKI, G. K. An evaluation of the postulates underlying the Babcock Deterioration Test. *Psychol. Rev.*, 1941, 48, 261-267.

GROUP PSYCHIATRIC SCREENING OF MILITARY RECRUITS[1]

By NELSON G. HANAWALT
NEW JERSEY COLLEGE FOR WOMEN
RUTGERS UNIVERSITY
AND
ROBERT A. KIDD
THE McMILLEN SANITARIUM
COLUMBUS, OHIO

PSYCHIATRIC screening of recruits entering the armed services during World War II was done, on the whole, by individual interview. The official Naval program called for a three minute interview for each recruit. Unfortunately this program was not feasible in many instances because of the limitation of psychiatric personnel. Consequently improvisations had to be made to meet the prevailing conditions. Group screening was devised under such circumstances.

At the time the group method was introduced the Training Center was receiving five thousand recruits a week. No more than three interviewers, including a psychologist, were available at any one time. Obviously three minutes could not be devoted to each recruit. The first recourse was to shorten the individual interview. This method proved unsatisfactory and inadequate. Paper and pencil tests, which had proved valuable in eliminating many recruits from the individual interview at other stations, could not be used because of administrative objections. Psychologists were assigned as interviewers[2], but since only one was available full time, this did not solve the problem. The idea of the group interview came from the reported success of group psychotherapy.

As the group method was developed it took various courses in the hands of the different operators. The size of the group interviewed at one time ranges from fifteen to twenty or more recruits. They are lined up in such a way that all can be closely observed during the interview. A short introductory speech is made in order to stress the importance of making the symptoms known at this time. Naturally the personality of the interviewer is important in establishing this rapport. It has been found that a firm and authorative approach, yet a kindly one, produces the best results. The important thing is to emphasize that you mean business and that you are looking out for the recruit's welfare.

[1]The material for this article was collected while the authors were on active duty in the United States Naval Reserve and stationed at a large Naval Training Station. The opinions expressed in this paper are those of the authors and are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large.

[2]In an unpublished study by one of the authors involving approximately the same sample as the present study, it was found that psychologists and psychiatrists were equally effective in picking out men in the initial brief interview who would be later discharged for neuropsychiatric reasons. There were substantial individual differences in this ability but background of psychology or psychiatry was not an important factor. Included in this study were four psychologists and six psychiatrists. Psychologists and psychiatrists were both represented in the two best and the two least successful in this screening process. The average scores of the two groups were about the same.

One method is to go through a group of two or more symptoms at a time and ask for confessions at the end of each group. By grouping the symptoms no one in the group knows specifically what any one member is admitting. In this grouping of symptoms the aim is to combine socially condoned symptoms with those which might bring forth embarassment, e.g. enuresis, nightmares, and somnambulism. After each group of questions the medical examination sheet of paper (which they all carry) is lifted of any recruit who admits to having any of the symptoms. This assures his remaining for further questioning after the others have gone. A typical interview by this method is outlined below.

The recruits are brought into the room and lined up by the corpsman. The interviewer has an opportunity to observe them walk in, respond to directions, and adjust to their places. This is important group behavior to observe. Hyperactivity, awkwardness, posture, tenseness, mannerisms, tics, timidity, selfconsciousness, and many other personality traits can be observed at this time. As soon as the recruits are adjusted the interviewer starts off with his introductory speech in an impressive tone of voice. It might run something like the following.

Boys this is the last doctor you are going to see before entering the navy. I want you to tell me if you have anything wrong with you, physically or mentally. Now is the time to make these things known. We are going to find out about these things anyway so you might as well make them known now. It will save you a lot of trouble.

Q. Let me see the hands of any who have ever been in the army, navy or marine corps on active duty. If any hands are shown they are questioned as to length of service and date of discharge. If they were in service less than the normal enlistment or if they were discharged after the beginning of the war they are held for further questioning. In this way undesirables, fraudulent enlistments, medical and inaptitude dischargees are singled out. This group is a fruitful one for further questioning.

Q. Has anyone been rejected for military service? Many who respond to this question are cases of poor vision, teeth, illness at the time of call or some other unimportant defect, so the reason for their rejection is ascertained. In this way a number are picked up who were rejected for immaturity, low intelligence, inadequacy, and any number of psychosomatic complaints. This too is a very fruitful group for further questioning.

Q. Has anyone been bothered with severe headaches, migraine headaches, dizzy spells, fits or convulsions, fainting spells, or blackouts?

Q. Has anyone had a fractured skull, brain concussion, or blow on the head which resulted in unconsciousness?

Q. Has anyone had nightmares, sleepwalking, falling out of bed, wetting the bed at night, or kidney trouble? (An enuretic will often admit to kidney trouble when he will not admit to wetting his bed.)

Q. Has anyone had recurrent stomach pains, ulcers of any kind, vomiting, or pains in his abdomen?

Q. Has anyone had asthma, spells of difficulty in breathing, or anything like that?

Q. Has anyone had rheumatism, arthritis, back pains, leg pains, or pains of any kind?

Q. Has anyone had a nervous heart, a jumpy heart, or heart burn?

Q. Has anyone changed jobs frequently or been arrested for anything aside from traffic violations?

Q. Has anyone had a nervous breakdown or nervousness of any kind?

At this point the importance of making symptoms known is again emphasized. Throughout the questioning it is well to keep adding words of encouragement for a confession. During the general questioning you have collected the medical sheets of some you wish to question further. Those who still have their papers pass before the desk to have their papers checked. As they step up to the desk they report their age and how far they went to school. Any question which applies to the individual recruit can be posed at this time. A man who has passed eighteen years of age can be

asked if he is married, if he has ever been fired from a job, or why he was deferred. A shy seventeen or eighteen year old might be asked if he has a girl friend, plays sports, etc. At this point a man who has not previously admitted anything might break down and confess. Those suspected of low intelligence or immaturity, stutterers and those with other speech defects, potential homosexuals, schizoid personalities, etc. are asked to remain. Most of these have been previously noted mentally because of telling behavior signs.

Usually out of a group of twenty there are three or four who are held for further questioning. Some of these are quickly disposed of because their symptoms are unimportant. Those suspected of low intelligence are asked a few questions and sent to the psychologist for further testing if it is warranted. To the remaining small group the individual technique is applied, but you have the advantage of much information and good rapport established before beginning the interview. A decision is made at this time whether these doubtful cases are to go to full duty, trial duty, or to the neuropsychiatric ward for further study. Those who are sent to trial duty are called in for another interview after three or four weeks. For this second interview a Company Commander's report is available on each one of them, and a Social History (if one was requested at the time of the original interview), and their reaction to a few weeks of actual training. This additional material makes possible a much better evaluation of their ultimate usefulness to the navy.

A variation of the above method is to ask the questions and have all stay who have any of the suggested symptoms or anything at all wrong with them. The important factor here is to get them to stay. In this method you ask for previous military service and rejections, as in the first method, before enumerating the symptoms. The list of questions has to be somewhat shortened in order that it is not beyond the memory span of the recruit. After the preliminaries are disposed of a speech something like the following is given.

> I want you to tell me if you have anything the matter with you. I want you to stay after the others have gone and tell me if you have severe and frequent headaches, dizzy spells, fits or convulsions, or fainting spells; if you walk in your sleep, fall out of bed, or wet the bed at night; if you have stomach trouble, back pains, or aches and pains of any kind; if you have ever had any nervous trouble of any kind, or if you have been arrested for anything excepting traffic violations. If you have any of these things, or if you have anything at all the matter with you, I want you to stay and tell me. It is very important to make these things known at this time. You are going to be sorry if you do not.

From this point the procedure is as in the first method. In this method, however, greater stress is put upon asking appropriate questions to the individual as he passes in front of the desk to have his paper checked. The group part takes less time than in the first method, consequently there is more time available for individual questions. This method has the advantage of allowing for more privacy for no one has to confess anything before the group. If he does want to talk, all he has to do is to say that he wants to stay. The first method has the advantage, however, of being more effective with a larger group.

Psychiatrists have to be sold on the group method, but nearly all of them, after giving it a try, agree that it is very effective where the pressure of time is too great for an adequate individual interview. A few have found their personalities unsuited to the group method and have continued to use the individual method. The two methods can be used side by side without conflict.

Since the group method was dropped for a period of time when additional

staff was available, it is possible to make a comparison of the effectiveness of the group method vs. the individual interview. In a series of five hundred consecutive cases coming before the Aptitude Board for discharge, eighty-seven per cent of them had been noted on the Receiving Line at the time of the original interview by the group method. During the time when only the individual interview method was used this percentage dropped to a little less than sixty per cent. This is no criticism of the individual method, but of the method as it had to be used under the particular conditions of the Receiving Line at this Station.

There are many advantages in using the group method. It is time saving. The recruit is under the observation of the Interviewer for a longer period of time. It affords an opportunity to observe the recruit in a group situation, which after all is very important in military life. It allows for asking each recruit more questions in a shorter period of time. Better rapport is established in the short time during which the recruit is available. In the individual interview you do not have time to emphasize the importance of "telling all." Also additional defects which are purely medical are confessed which can be noted for observation or entry into the recruit's medical record.

The chief criticism of the method is the lack of privacy for the interview. This criticism is far less important in the military situation than one would suspect from experience in private practice. The recruit has been pushed around, deprived of sleep, shouted at, probed here and there, deprived of his clothes and hair, etc. before coming to the psychiatrist. His resistance is pretty well broken down. For instance, enuresis is one thing that one might suspect would cause embarrassment to admit in the group situation, yet it has been found that it is much more readily admitted in this situation than in the individual interview. Also due respect for privacy can be allowed by the group method. The second method described above emphasizes privacy, but it has been found to be no more effective than the first method which calls for open confession.

It has been said that the group method is less thorough than the individual interview. This criticism is made by those who are thinking of the individual interview as used under favorable circumstances. Certainly the individual interview applied under favorable circumstances would be superior to the group method as it has been used here under the pressure of limited time per recruit. No check has been made of the effectiveness of the two methods where both have been used under favorable conditions. However the results of the comparison reported above demonstrate that where the time per recruit is greatly curtailed, the group method is far superior. The group method was introduced as an emergency measure but it is the opinion of the writers that it has many advantages over the individual method in military screening even though enough time is available for a three minute interview for each recruit. In our experience one of the chief problems in military screening is to dispose of those who are suitable with dispatch and find time for a thorough study of the borderline material. In this respect, the group method has obvious advantages, with its individual contact with the recruit, over the paper and pencil method which has sometimes been used for this purpose.

Received February 16, 1949.

THE RELATIONSHIP OF INTELLIGENCE AND YEARS OF SCHOOLING WHEN LITERACY IS HELD CONSTANT

BY WILLIAM D. ALTUS
SANTA BARBARA COLLEGE
UNIVERSITY OF CALIFORNIA

IF one were asked why some adult males cannot read at all, or at best can read so little that they are functionally illiterate, he would probably reply, as the writer once did, that the two most logical reasons would be (1) lack of verbal intelligence or (2) lack of opportunity to learn. It does take a certain amount of intelligence to learn to read and if an adult is a functional illiterate he is usually not very bright. The average IQ on the Army Wechsler for more than 15,000 Army illiterates during the late war was about 67. Even among these illiterates of restricted intellect, there was a significant relation between their Wechsler scores and their amount of literacy as defined by Gray's Oral Reading Test. The r for 250 cases was .37 between these two variables. Such a correlation does not appear to be high but when one remembers the attenuation of range in Wechsler and in literacy scores, such an r is seen to be considerably more significant than it would appear simply from its size.

Since the more intelligent of the functionally illiterate soldiers could read somewhat better than the less intelligent, the writer was curious as to what the order of relationship would be between years of schooling in a public school and Wechsler IQ if literacy and certain other distracting variables were held constant. The population of this research, then, consisted of old-line, native-born American Whites only, who had been born and brought up in Washington, Oregon and California. In this manner, race, linguistic influences (absence of bi-linguals) and regional differences in quality of schools were fairly well controlled. The records of the Ninth Service Command Special Training Center were then combed for those trainee soldiers who fitted these categories; those cards, finally obtained, were further sifted for those who scored exactly zero, four, ten or sixteen on Gray's Oral Reading Paragraphs. A score of zero would be equivalent to no schooling; a score of four to end-of-first-year schooling; a score of ten to second-year schooling; and a score of sixteen to third-year schooling. The resulting records permitted the calculation of Pearson product-moment coefficients of correlation between years of schooling for each of the Gray-defined literacy groups and Wechsler IQ's. The data are presented in Table I.

TABLE I
RELATION BETWEEN IQ AND YEARS OF SCHOOLING WHEN READING ABILITY IS HELD CONSTANT

Score on Gray's Paragraphs	Mean Wechsler IQ	Mean years of schooling	N	r, school vs. Wechsler
Zero	65.75	6.21	85	+.20
Four	71.10	8.10	21	+.38
Ten	72.60	7.25	17	+.19
Sixteen	70.85	7.68	22	—.07

It will be noted in this table that there is a general tendency for the more liter-

ate trainees to have a higher mean Wechsler IQ. For instance, those who read at end-of-second-year level had a mean IQ of 72.60, which is approximately seven points higher than the mean IQ of 65.75 for the zero readers. But within groups it will be noted that the correspondence of opportunity to obtain the rudiments of literacy as defined by the years of schooling and the intelligence quotient goes, three times out of four, in a direction opposite from that which would be expected. None of the coefficients, it is true, quite reach significance, though the first two shown are very close to the .05 level of significance. This lack of significance is not, however, nearly so surprising as is the direction of the relationship for three of the four r's reported. It would appear that there is a slight tendency, when literacy and certain other distracting variables are held constant, for "brighter" individuals to have attended school for a longer period of time.

On the face of it, then, it would seem that there is a relationship between intelligence and persistence in attendance in the elementary schools—at least, it can be said that three of the four r's definitely point in that direction. The relatively duller illiterate had stayed in school, on the average, a shorter period of time, yet had achieved the same degree of literacy (or illiteracy if one wishes to look at the obverse side of the medal) as his relatively "brighter" brother illiterate. Since the data are so much at variance with what the writer expected to find, the subtitle of this short paper should be "Some data in search of an explanation." Many possible explanations have been considered and rejected. Can it be that simply remaining in school increases one's intelligence slightly, even though the putative *sine qua non* of instruction, namely literacy, is not present? Or, since none of the r's reach even the .05 level of significance, are the data simply to be regarded as interesting bits of information of no significance whatever?

Received January 11, 1949.

LIFETIME WORRY PATTERNS OF AMERICAN PSYCHOLOGISTS[1]

By WILLARD A. KERR, HARRY L. NEWMAN, and ALFRED R. SADEWIC

ILLINOIS INSTITUTE OF TECHNOLOGY

As the first of a series of projected studies of the "lifetime" worry patterns of the members of specific American occupational groups, this research employed the voluntary, anonymous questionnaire technique.

A letter-questionnaire requesting cooperation and listing twelve potential fields of worry was prepared and mailed with attached postpaid answer - reply card to a random sample of 250 successful American psychologists who were born before 1903. "Successful" for the purpose of this study is defined as being employed in a psychological capacity, and, with a few exceptions, holding fellowship status in the American Psychological Association. Ages of the 103 respondents ranged from 46 to 81, the median age being 54. Each psychologist was requested to state the age or age level at which any or each of the potential worries had been most of a problem to him.

[1]Grateful acknowledgment must be made to Dr. Dael Wolfle for constructive criticism in interpreting the results of this study.

Inspection of Table I reveals economic worry to be the most frequently reported

TABLE I

STATISTICAL SUMMARY OF REPORTED WORRY EXPERIENCE OF 103 MALE PSYCHOLOGISTS BORN BEFORE 1903

		Per cent Expressing Worry	Median Age at Time of Worry	Mean Age at Time of Worry	Inter-Quartile Range
1.	Appearance	52	20	20.8	16-31
2.	Sexual morality	54	23	24.5	18-30
3.	Lack of confidence of making a good personal impression in meeting people	65	26	24.7	19-39
4.	Religious or philosophical convictions	88	26	28.2	21-41
5.	Economic worries	91	30	32.0	23-38
6.	Lack of confidence of business success	63	31	33.3	26-38
7.	Job security	75	32	34.6	26-38
8.	Peace of mind	67	35	34.8	26-47
9.	Willing to return to job at less pay	50	36	40.3	32-46
10.	Political convictions	82	38	37.8	29-48
11.	Health	57	38	40.5	28-50
12.	Giving up important hopes or ambitions	56	40	38.5	28-45
13.	Marital difficulties	42	41	40.2	34-48

of the twelve potential worries. Nine out of ten psychologists state a specific age or age level of economic worry, the median age of such worry being 30.

Since economic worries appear to reach their maximum at age 30, a first hypothesis that seems reasonable to account for this result is that a maximum number of respondents just happened to be age 30 during economic depressions. In order to evaluate this hypothesis, Figure 1 was developed.

When the respondents were divided into four age groups (see Figure 1) approximately one fourth were found to be in each group. The frequency distributions of occurrence of economic worry of the four age groups were then plotted and the years of national economic depression were also indicated. Inspection of Figure 1 reveals that the median or modal ages of economic worry do not occur in economic depression periods much, if any, more frequently than would be expected by chance. Psychologists aged 45-49 for example, reported a peak economic worry load at age 31 which was in the depression year of 1932, but those psychologists aged 50-54 also reported a peak economic worry load at age 31 and that was the relatively prosperous year of 1927. Examination of the areas under the curves in Figure 1 shows some tendency toward greater economic worrying during depressions, but the tendency toward peaking of economic worry between ages 29-34 persists regardless of the business cycle.

Actually two other of the twelve worries studied seem psychologically relevant to economic worry— job security and lack of confidence of personal business success. This belief is supported

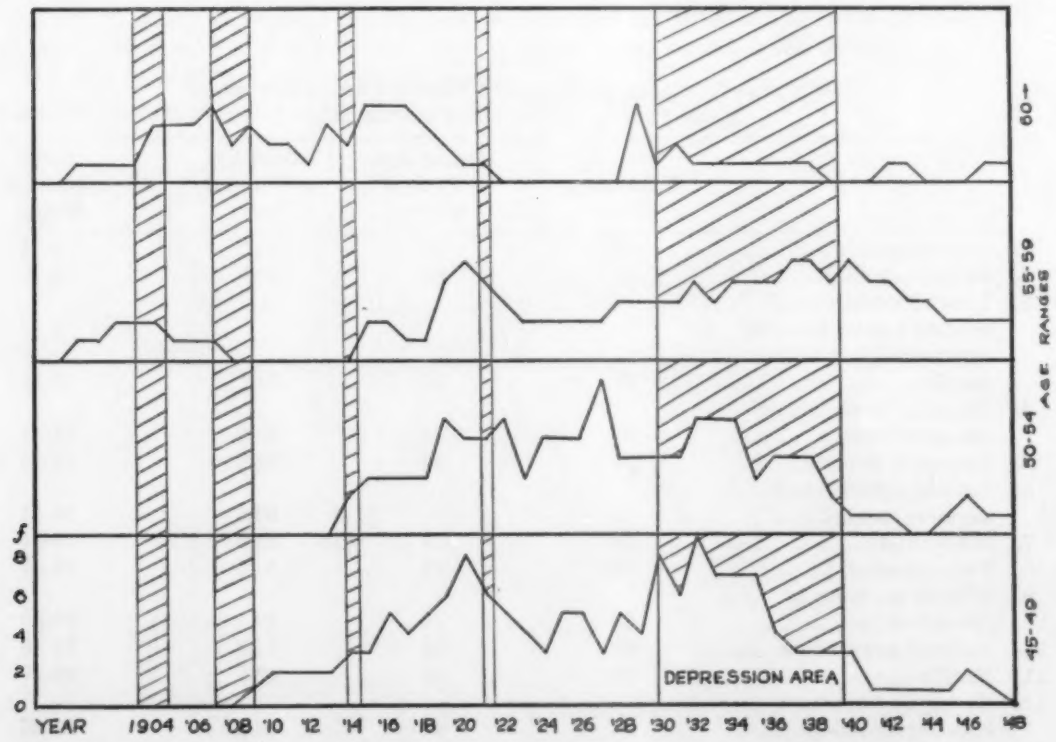

FIG. 1. Economic worries of 103 male psychologists born before 1903.

by the obtained data of Figure 2, which reveals the modal point of worry for all three worries at age 30; also, as shown in Table I, the quartile deviations of these worries are practically identical.

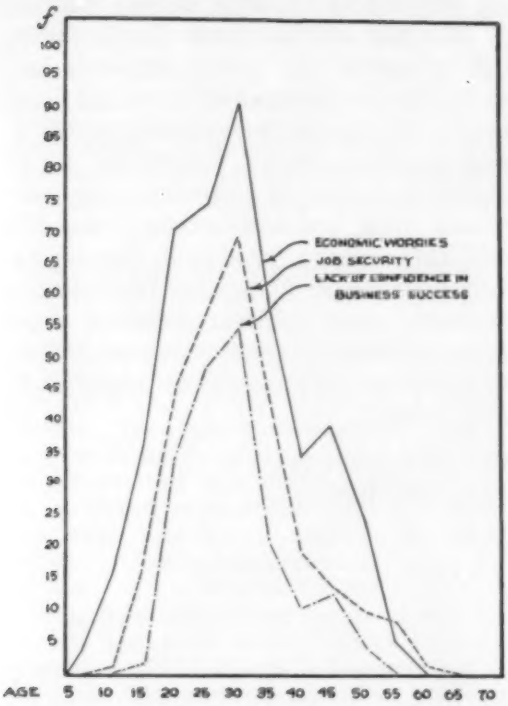

FIG. 2. Vocational-economic worries of 103 male psychologists born before 1903.

The central tendency data of Table I provide a probable chronological sequence of twelve of the major worries which seem likely to arise in the lifetime of typical American psychologists. This sequence begins with worry about personal appearance maximizing at age 20, overlapping markedly with worries related to sexual morality which peak at age 23.

Proceeding from his early twenties amidst concern over personal appearance and sexual morality, the psychologist at age 26 apparently becomes more worried than at any other time over lack of confidence in making a good impression in meeting people and over religious convictions. At age 30, a notable change in the worry load seems to occur with peak emphasis on economic worry —which incidentally is the same age as the mean age of marriage of the respondents. At age 31, maximum lack of confidence of personal business success is encountered, and age 32 finds greatest worry over job security. General worry over peace of mind lacks any notable age peak, but the median age of realization of lack of peace of mind is 35.

At age 36 the maximum tendency apparently is felt to take a job of less pay. This would seem to indicate a psychological breaking away from the economic worries of the preceding six-year period. The shift from personal economic worries is further reflected at age 38 and thereafter by a continuing increase in emphasis on political convictions.

Forty-two per cent of the respondents indicated marital difficulties as being sources of worry; the median age for these worries is reported as 41.

Two variables yielded interesting bimodal distributions. "Giving up ambitions" reaches a small peak at age 22 and a high plateau at ages 39 - 45. Health as a worry reaches a high plateau at ages 30-35 and another plateau proceeding on from age 45.

Reliability of these worry patterns at the oldest age levels undoubtedly is least because of diminishing numbers of cases.

SUMMARY

Many difficulties are inherent in this type of research, especially the unknown effects of the memory factor, the biases of the 40 per cent who cooperated, and the possible tendency to emphasize past rather than present worries, therefore, all conclusions indicated by statistical analysis must be regarded as highly tentative; nevertheless, within the limitations of the study, certain conspicuous

trends in age-worry pattern characteristics are apparent.

1. Of the twelve potential worries reported by 103 older psychologists, economic worry was reported with greatest frequency.

2. Economic worries apparently reach their peak near age 30, this tendency persisting in periods of both economic depression and prosperity.

3. As reviewed in perspective, the worry characteristics of the typical older psychologist responding with reference to age of maximum worry fall into the following sequence: age 20 — personal appearance, age 22—giving up ambitions, age 23—sexual morality, age 26—making a good impression in meeting people and concern over religious or philosophical convictions, age 30 — economic problems, age 31—personal business success, age 32—job security, age 33—health, age 38 and thereafter—political convictions, age 41—marital difficulties, age 42—giving up ambitions, age 45 and thereafter—health.

4. When aged about 18-26, the typical respondent reported chiefly worries of idealism and personal development, but by age 30 the worry emphasis has shifted to problems of survival and hard reality. By age 38 the economic survival worries have eased somewhat as is shown by maximal and continuing emphasis upon political convictions with political concern gradually increasing into later years. This later life relative emphasis upon political problems indicates a practical combination of earlier idealism and subsequent experience with reality.

Received July 5, 1949.
Early publication.

New Books and Tests

BOOKS

AARON, BETTY. *A manual for analysis of the Thematic Apperception Test.* Berkeley, Calif.: Willis E. Berg, 1512 Spruce St., 1949. Pp. xiii + 163. $3.50.

Since its first appearance fourteen years ago, the TAT has remained an elusive technique, despite its wide use. Although Murray originally proposed a somewhat quantifiable scoring in terms of needs and presses, many users have preferred a "global" or "intuitive" interpretation. The apparent unsatisfactoriness of the original scoring might be due to either of two defects: that it is faulty in principle, or that it is insufficiently elaborated and developed. Miss Aaron's work supports the latter hypothesis. Her manual presents a further development of the Murray-Sanford scoring of content, which may be checked for scoring reliability, and which permits the quantitative comparison of individuals and groups, a first step toward experimental validation. Although complicated and not rapid, the scoring technique is clear enough to be mastered, and is at least practicable for research. As yet, the research procedures do not permit statistical tests of significance, which may lead to unjustified conclusions. For example, in one of Miss Aaron's sample research studies, she concludes that "a general decrease in expression is found at the pre-puberty stage," on the basis of a lowered output in three out of four groups of children 11 to 12 years old, numbering 4 or 5 in each group. Any step of true research in the projective field is encouraging, as is Miss Aaron's study, but much of the path of progress lies ahead.

BRANHAM, VERNON C. AND KUTASH, SAMUEL B. *Encyclopedia of criminology.* New York: Philosophical Library, 1949. Pp. xxxvii + 527. $12.00.

While "encyclopedia" is perhaps an ambitious title for a volume of some five hundred pages, the work is not without value. Sixty-one contributors, from psychiatrists to police chiefs, wrote its 118 articles. Each reader will profit most from the essays outside his field of competence: psychologists may want to know something about parole, and lawyers about mental deficiency. As in most such works, the quality of articles varies widely, some are too technical, some are "written down" to the point of cuteness, many seem to strike about the right level.

CAVAN, RUTH S., BURGESS, ERNEST W., HAVIGHURST, ROBERT J., AND GOLDHAMER, HERBERT. *Personal adjustment in old age.* Chicago: Science Research Associates, 1949. Pp. xiii +204. $2.95.

The first report of a ten-year research program on the adjustments of old age gives data based on a survey of about three thousand people over sixty years old, with more intensive studies of smaller groups. The major research instrument was a specially developed questionnaire on attitudes and interests (reviewed below), but case studies were also made, as well as several types of observations and ratings of the subjects by others. The mass of detailed findings, presented in descriptions and many tables, does not yield readily to sweeping generalizations. Some results, as the increase in illnesses with age, seem obvious; others, as the general increase in feeling of economic security with advancing age, are not so obvious but just as true. The study has great value for all concerned with problems of aging, but has only begun to penetrate its large field.

DENNIS, WAYNE. *Readings in general psychology.* New York: Prentice-Hall, 1949. Pp. xi + 525. $5.00.

A first-hand acquaintance with a well-se-

Note: Some reviews in this issue were prepared by the Associate Editors, who may be identified by their initials. Unsigned reviews are by the Editor.—L. F. S.

lected sample of historic writings and classic experiments may lead the beginning student in psychology to seek original sources and not to depend on secondary accounts alone. Dennis has assembled 70 excerpts from books and articles, ranging from daVinci (1500) to N. E. Miller (1948), into an effective teaching device. The penetration of clinical interests into the introductory course is seen in the large number of selections in the areas of emotion, motivation and personality. Incidentally, some students of clinical psychology, far beyond the introductory level, might read these articles with profit.

GILBERT, JEANNE G. AND WEITZ, ROBERT D. *Psychology for the profession of nursing.* New York: Ronald, 1949. Pp. x + 275. $3.00.

This introductory text for students of nursing, covering the entire range of behavior from normal to abnormal in only two hundred fifty pages, is of necessity brief and superficial in its treatment of specific topics. It is poorest in its handling of such concepts as sensation, perception, learning, and thought; and at its best in dealing with such practical problems as the psychological aspects of caring for the aged. In view of its brevity, its passing affirmation of faith in mental telepathy seems unnecessary.—*W. A. H.*

GRINKER, ROY R. AND BUCY, PAUL C. *Neurology.* (4th Ed.) Springfield, Ill.: Charles C. Thomas, 1949. Pp. ix + 1138. $12.50.

The fourth edition of this important text and reference book in neurology has been revised completely, mainly by its junior author. The book stays entirely within the field of its superb competence, the organic disorders of the nervous system. All for the best, it lacks the usually superficial sections on functional psychiatry and clinical psychology that some other neurology books have. A brief paragraph acknowledges the value in organic diagnosis of tests and projective methods which "require the auxiliary services of a trained psychologist." With its authoritative and up-to-date coverage, its clear writing, and its exhaustive 112-page index, the volume is an efficient reference resource for the clinical psychologist who works with neurologists.

HOVLAND, CARL I., LUMSDAINE, ARTHUR A., AND SHEFFIELD, FRED S. *Experiments on mass communication.* Studies in social psychology in World War II, Vol. III. Princeton, N. J.: Princeton Univ. Press, 1949. Pp. x + 345. $5.00.

The experiments reported here evaluated films and radio scripts used for indoctrination and instruction in the Army. The large samples and realistic conditions permitted experiments that would be difficult to perform in civilian life. On the other hand, the military preoccupation with immediate and practical ends sometimes hampered basic research and the independent manipulation of all pertinent experimental variables. The volume is a model of presentation, technical enough for specialists but comprehensible to generally educated laymen.

HUMPHREY, GEORGE *Directed thinking.* New York: Dodd, Mead, 1948. Pp. ix+229. $3.50.

The head of psychology at Oxford University brings the experimental psychology of the thought processes together in a vividly readable book for laymen. As he indicates in a "postscript for psychologists," the author draws on many experiments, often not cited specifically, and uses his anecdotes as illustrations, never as proofs, of the findings. You will enjoy this book. If you are something of an old-timer, you will remember when you first read Dewey's *How we think* twenty or thirty years ago. There has scarcely been another such book in the interval.

ISSACS, SUSAN. *Childhood and after.* New York: International Universities Press, 1949. Pp. viii + 245. $4.50.

This is a collection of thirteen papers and addresses published or delivered by Susan Isaacs between 1928 and 1945, ten of them previously published. Subjects vary from "The mental hygiene of the pre-school child" to "A special mechanism in a schizoid boy." In all the papers, the orthodox Freudian point of view prevails, although the treatment varies in intensity in accordance with the audience or readers for whom the papers were intended —in some cases the lay public, and in others, professional psychoanalysts. In spite of its unevenness, this is an important volume for the clinical psychologist, particularly if he deals with children. Those not psychoanalytically oriented may gag at some of the dynamics described, but that should not deter one from profiting from some of the best discussions of the development of children's personalities found in the psychological literature. —*M. K.*

JONES, ERNEST *Papers on psychoanalysis.* (5th Ed.) Baltimore: Williams & Wilkins, 1948. Pp. vii+504. $8.50.

The fifth, which the author says "should be positively the final," edition of Jones's papers brings together 27 selections spanning the period from 1909 to 1947. Fourteen papers that appeared in the preceding edition have been omitted and nine newer ones have been added. The viewpoint of Freud's missionary to the English, and therefore to a great extent to the Americans, has changed little with time. His preeminence as a clear interpreter and developer of orthodox psychoanalysis remains unchallenged. Many moderns will disagree with his continued support of an inherited and fixed process of personality development, which rejects the weight of cultural influence.

KANNER, LEO *A miniature textbook of feeblemindedness.* Child Care Monographs No. 1. New York: Child Care Publications, 1949. Pp. 33. $1.25.

A short readable monograph that should be seen by physicians, educators, and quite a few psychologists too, who have retained the conceptions of feeblemindedness current a generation ago. Kanner traces the recent history that has separated many separate types from the previously undifferentiated mass of mental deficiency. He groups the feebleminded into three groups, the absolute, relative, and apparent, whose treatments and prognosis differ. The genetic, cultural, material, physical, educational and emotional determinants are discussed briefly. There are several short case studies, and 71 references.

LAPIERE, RICHARD T. AND FARNSWORTH, PAUL R. *Social psychology.* (3rd Ed.) New York: McGraw-Hill, 1949. Pp. xiii + 626. $4.50.

The new third edition of this well-known text does not depart from the essential plan of the preceding one. New research studies have been added, and there are three entirely new chapters in the areas of socialization, personality, and the social control of deviations. There is much of value for the background of clinical psychologists, not all of which is represented in every book in social psychology, e.g. the discussion of personality "types" in terms of the need of stereotypes in place of a really intimate knowledge of people. Its mature, eclectic and readable character make this book one of the best choices of a text in its area.

MATHEWSON, ROBERT H. *Guidance policy and practice.* New York: Harper, 1949. Pp. xiv+294. $3.00.

The author develops a philosophy of guidance policy from psychological, social and educational bases. In the main, he favors an organismic view of personality and a self-directive approach to counseling. In this framework, guidance is seen as a continuing experience helping normal persons toward self-realization, not as a remedial service to be invoked only in crises or for deviants.

MOON, GEORGE R. *How to become a doctor.* Philadelphia: Blakiston, 1949. Pp. ix + 131. $2.00.

Intended primarily for the student who seeks to enter medicine and for the information of guidance counselors, this small book has interesting incidental values for psychologists concerned with professional training. Written by the Examiner and Recorder of the University of Illinois College of Medicine, it gives a very detailed picture of medical school admission practices, with samples of forms used. The descriptions of the medical curriculum, the accreditation of schools, and the licensing of physicians throw light on analogous problems of our profession. Dentistry, pharmacy and a number of other professions allied to medicine are also discussed, but clinical psychology is not included.

MURSELL, JAMES L. *Psychological testing.* (2nd Ed.) New York: Longmans, Green, 1949. Pp. xvi + 488. $4.00.

The revision, coming less than two years after the publication of the first edition (see *J. consult. Psychol.*, 1948, 12, 60), corrects a number of the shortcomings of a generally well-organized and useful book. There are increased references to research studies, and a number of important recent tests have been added.

OLSON, WILLARD C. *Child development.* Boston: Heath, 1949. Pp. xiii + 417. $4.00.

A scholarly and comprehensive discussion of child growth and development during the early and middle years of childhood. The book is planned to meet the needs of those working with children but especially educators. It reviews the research in various areas of child development — physical, motor, play, mental and emotional growth, and school achievement. In spite of the fact that the reviews are group-

ed under various topics, there is a constant emphasis upon the whole personality and the interplay of forces that affect the growth and development of the child as an individual. The book attempts to weld the mass of data and research in these various areas into a real working philosophy that educators can understand and use. While constantly emphasizing the whole personality it reviews the significant studies in the areas of individual and group relationships. Dr. Olson points out those areas where he has found research limited and further study needed to clarify the findings. This comprehensive digest of research studies will enable the educator to re-evaluate his own working philosophy and reinterpret educational procedures in the light of present findings.—*B. M. L.*

PIÉRON, HENRI. *Traité de psychologie appliquée. Livre I. La psychologie différentielle.* Paris: Presses Universitaires de France, 1949. Pp. viii + 122. 240 fr.

Piéron's treatise on individual differences is mainly a digest and summary of research, with little space devoted to abstract theorizing. There are three chapters, on heredity, aptitudes, and the problem of types. The extensive bibliographies will look familiar, as a large proportion of citations are from American sources. In form and purpose, the volume resembles our *Handbooks*, but it covers the applied field, in which we still lack a comprehensive compilation. Forthcoming volumes, under Piéron's editorship, will deal with the methods of applied psychology, applications of aptitude measurement, educational method, human management, mental hygiene, and social fields of application.

SUPER, DONALD E. *Appraising vocational fitness, by means of psychological tests.* New York: Harper, 1949. Pp. xxi + 727. $6.00.

No previous work has surveyed the application of psychological tests to vocational guidance and selection so comprehensively. After orienting chapters on the general problems of the field and on test construction and administration, Super gets to the core of his material in fourteen chapters that deal with the research evidence on particular tests. One handicap of his approach is that the older tests with a large research literature receive the most space, not always in proportion to their merit shown by the evidence. The emphasis on experimental findings excludes vague generalities, however, and often points to gaps in our knowledge that should stimulate research. The book concludes with chapters on the use of tests in counseling, with illustrative case studies. The bibliography has 955 entries. Psychologists who use tests in vocational counseling will find the book valuable for their own reference use as well as indispensible for their advanced students.

THOMAS, WARD. *Stranger in the land.* Boston: Houghton Mifflin, 1949. Pp. 373. $3.50.

A novel of an anxiety-ridden homosexual and a blackmailing psychopath, which rings fairly true as a clinical history.

WALKER, KENNETH, AND STRAUSS, ERIC B. *Sexual disorders in the male.* (3rd Ed.) Baltimore: Williams & Wilkins, 1948. Pp. xiii+260. $3.50.

Two British physicians prepared this objective but relatively nontechnical account of the common difficulties of the male sexual function. They recognize clearly the psychogenic origin of most such disorders, and constantly illustrate the use of an eclectic but by no means superficial psychotherapy in treating them. If the authors communicate to their patients the understanding and calm attitude that their writing shows, they are surely helpful therapists. Readers may well assimilate and adopt their attitudes.

Poliomyelitis. Papers and discussions presented at the First International Poliomyelitis Conference. Published for the National Foundation for Infantile Paralysis, Inc. Philadelphia: Lippincott, 1949. Pp. xi+360.

While the conference that this volume reports was mainly concerned with the medical aspects of poliomyelitis, several of the panel discussions also showed a recognition of the problems of the infantile paralysis patient as a person. Several psychiatrists and two clinical psychologists participated as panel members.

Social work as human relations. Anniversary Papers of the New York School of Social Work and the Community Service Society of New York. New York: Columbia Univ. Press, 1949. Pp. viii + 288. $3.75.

The family in a democratic society. Anniversary Papers of the Community Service Society of New York. New York: Columbia

Univ. Press, 1949. Pp. viii+287. $3.75.

These companion volumes commemorate the one hundredth anniversary of the Community Service Society and the fiftieth of the New York School of Social Work. They show how closely psychology and social work have drawn together in recent years as each has moved toward the other's position in emphasis on theory and service. Even much of their research becomes indistinguishable, as social and clinical psychology acquire social work's broad cultural frame of reference, and social work acquires the former's refinements of technique. *Social work as human relations* contains three parts: theory and techniques, including much that is psychological; professional training, which should be read by all of us concerned with training in psychology; and vistas in human relations, a presentation of the broader world view. *The family* is divided into two sections, on sociological-psychological studies, and on health and nutrition. The two volumes give a competent vignette of what progressive social work is thinking today.

TESTS

ARTHUR, GRACE *Arthur Adaptation of the Leiter International Performance Scale.* Ages 4-8. Individual test. 1 form. Test material ($55.00); carrying case ($7.50); record cards ($3.50 per 300); temporary manual ($1.00), revision in preparation. Washington, D. C.: Psychological Service Center, 822 Eighteenth St., N. W., 1949.

Arthur has standardized a children's form of the Leiter scale, using the tests of year levels II through X and XII, with a few items rearranged. The interpretation uses Arthur's point scale score method. Information on standardization procedures and data on reliability and validity have not yet been published.

BURGESS, ERNEST W., CAVAN, RUTH S., AND HAVIGHURST, ROBERT J. *Your activities and attitudes.* Adults, ages 60+. 1 form. (60-90) min. Questionnaire blank ($2.75 per 25) with manual, pp. 4, and keys. Chicago: Science Research Associates, 1949.

The first standardized inventory intended specifically for elderly adults was prepared for the research study on personal adjustment in old age (reviewed above). It consists of three parts, an adult activities inventory measuring reported social participations, an attitude inventory, and questions on background information. Reliability of the activities inventory is shown by 82 per cent agreement of replies on retest, and of the attitude scale by correlations of .72 on retest and of .95 by the corrected split-half method. Validities estimated from correlations with various ratings by others range from .53 to .65, and the activities and attitude scores correlate .78. Uses are suggested for research, and for practical applications in group work, social work, personnel work, and individual counseling.

CASSEL, ROBERT H. *Vineland Adaptation of the Oseretsky Tests.* Ages 4-16. Individual test. 1 form. (45) min. Manual: Training School Bulletin, Monograph Supplement No. 1, pp. 32. Vineland, N. J.: The Training School, 1949.

The Vineland adaptation simplifies the materials needed for the Oseretsky tests of motor ability, and makes the tasks more specific and objective. The 45 tasks, plus 2 alternates, are in the areas of general static, general dynamic and dynamic manual coordination, speed, and simultaneous voluntary movement. Tests of asynkinesia, included in the original version, are omitted because the tasks were insufficiently objective to permit reliable scoring. Although the monograph contains a report of the performances of 54 feebleminded subjects, the test is regarded as not yet standardized.

GILLILAND, A. R. *Northwestern Intelligence Tests. Test A. Test for Infants 4-12 Weeks Old.* Ages 4-12 weeks. Individual test. 1 form. Materials, 12 items, may be assembled or obtained from author. Record blank ($2.00 per 25), with manual, pp. 15; specimen set (40¢). Boston: Houghton Mifflin, 1949.

Because of the need, especially of adoption agencies, for a test for very young infants, a scale of 40 items was developed drawn in part from the work of Gesell, Cattell and Buehler and in part from original observations. The observations emphasize responses showing adaptation to the physical and social environment rather than growth or maturation. All items show developmental progress in the per cent of infants passing with each succeding week of age. Reliability is .84 by corrected split-half, and .87 by test-retest. Norms, in terms of MA and IQ, are based on the testing of 1100 infants. There is some evidence of good agreement with other infant tests.

LEITER, RUSSELL G. *Leiter International Performance Scale, 1948 Revision.* Ages 4-adult. Individual test. 1 form. Test material ($82.50); carrying case ($7.50); record cards ($3.50 per 300); manual ($1.00), revision pending. Washington, D. C.: Psychological Service Center, 822 Eighteenth St., N. W., 1948.

The revision of the Leiter scale uses the same tests as the Arthur adaptation for years II through XII (*see above*), together with tests at the XIV, XVI and XVIII levels. A new adjustable frame to hold the test strips is provided in place of the six frames required for the 1940 version of the test, and the materials are now portable. Although the scale seems to be a stimulating and original contribution to the measurement of ability with nonlanguage tasks, full clinical acceptance still awaits the publication of data concerning its standardization.

MANSON, MORSE P. *The Manson Evaluation.* Adults. 1 form. Questionnaire form ($2.50 per 25, $9.00 per 100), with manual, pp. 2; specimen set (25¢). Beverly Hills, Calif.: Western Psychological Services, 1948.

The "Manson Evaluation" is disguised title of a questionnaire standardized to discriminate alcoholic personalities from nonalcoholics. Its 72 items were selected from 470 used in trial forms, and are mainly discriminative at the 1 per cent level. The critical score (for men) detects 79 per cent of alcoholics, with 21 per cent of false positives. Kuder-Richardson reliability is .94. The scoring permits a profile analysis of seven item clusters characteristic of alcoholics: anxiety, depressive fluctuations, emotional sensitivity, resentfulness, incompleteness, aloneness, and disturbed interpresonal relations.

PARTINGTON, JOHN E. AND LEITER, RUSSELL G. *Partington Pathways Test.* Adult males. 1 form (2 parts). (10) min. Test blanks, 2 practice sheets, 2 test sheets ($1.50 per 100 of each form), with manual, pp. 20; record card ($3.00 per 100). Washington, D. C.: Psychological Service Center, 822 Eighteenth St., N. W., 1949.

The Pathways Test is a performance task using printed forms. In Part I the subject draws a line to irregularly placed circles numbered 1, 2, 3, etc., in sequence, and in Part II to similar circles in the order 1, A, 2, B, 3, C, etc. Although an r of .68 with the 1937 Stanford Binet is reported for 256 adult males, the test is not intended to be a comprehensive measure of intelligence. The author regards it as a measure of a complex of speed, eye-hand coordination, alertness and distributed attention. Norms are given in terms of MA and IQ, for adult males only.

PSYCHOLOGICAL BULLETIN

July
No. 4
1947

THE PRESENT STATUS OF PSYCHOTHERAPEUTIC COUNSELING

BY

WILLIAM U. SNYDER

This popular review is the July, 1947 issue of the Psychological Bulletin

Pp. 90 $1.25

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 MASSACHUSETTS AVENUE, N. W.
WASHINGTON 5, D. C.

From the Best Seller List of The American Psychological Association

FOR FALL 1949

KUNST, MARY S.—A study of thumb-and finger-sucking in infants. *Psychological Monographs: General and Applied*, 1948, #290. $1.25.

ORLANSKY, HAROLD.—Infant care and personality. *Psychological Bulletin*, 1949, January (#1). $1.25.

SNYDER, WM. U. — The present status of psychotherapeutic counseling. *Psychological Bulletin*, 1947, July (#4). $1.25.

EVANS, JEAN. — Johnny Rocco. *Reprinted from the Journal of Abnormal and Social Psychology*. $.25 for single copy; $10.00 for 50 copies.

WOLFLE, HELEN M.—1949 *Directory, American Psychological Association*. $2.00

AVAILABLE BACK ISSUES

JOURNAL OF CONSULTING PSYCHOLOGY

The issues listed below are available for sale

YEAR	VOL.	NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
		JAN. FEB.	MAR. APR.	MAY JUNE	JULY AUG.	SEPT. OCT.	NOV. DEC.		
1937	1	1	2	3	4	5	6	$.60	$3.00
1938	2	1	2	3	4	5	6	$.60	$3.00
1939	3	1	2	3	4	5	6	$.60	$3.00
1940	4	1	2	3	4	5	6	$.60	$3.00
1941	5	1	2	3	4	5	6	$.60	$3.00
1942	6	1	–	3	–	–	6	$.60	$1.80
1943	7	–	2	3	4	5	6	$.60	$3.00
1944	8	1	2	3	4	5	6	$.60	$3.00
1945	9	–	2	3	4	5	6	$.60	$3.00
1946	10	1	2	3	4	5	6	$.60	$3.00
1947	11	–	–	3	4	5	–	$.60	$1.80
1948	12	1	2	3	4	5	6	$1.00	$5.00
1949	13	By Subscription, $5.00						$1.00	

Prices for Volumes 1 through 12 $35.60

Table based on inventory of January 3, 1949

Information about the Journal of Consulting Psychology: all numbers but 8 are still available.

Information about prices: the Journal of Consulting Psychology had the uniform price of $3.00 per volume and $.60 per issue from 1937 to 1947. It is now priced at $5.00 per volume and $1.00 per issue. For foreign orders, $.25 per volume should be added. Address orders to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 Massachusetts Ave. N. W.
Washington 5, D. C.

Distinctive McGRAW-HILL *Books*

MENTAL HYGIENE IN PUBLIC HEALTH

By PAUL V. LEMKAU, The Johns Hopkins University. *McGraw-Hill Series in Health Science.* 396 pages, $4.50.

This is the first book of its kind in preventive psychiatry to treat the problem of mental health in terms of public health action. Dr. Lemkau traces the development of the personality, illustrating at each stage and change in development the possibilities for modification in the direction of better mental health. An effective basis for setting up constructive programs is presented.

PERSONALITY MALADJUSTMENTS AND MENTAL HYGIENE New 2nd Edition.

By J. E. WALLACE WALLIN, Upsala College. *McGraw-Hill Publications in Psychology.* 568 pages, $5.00.

An extensive revision and reorganization of a well-known book. Some topics have been expanded, new material has been added, and the text has been generally brought up to date. Whole chapters are now devoted to a discussion of the educational elements of the mental hygiene program, psychotherapeutic methods, and the practical mental hygiene treatment of inferiority feelings, compensatory maladjustments, and fears. Many of the old case histories have been eliminated in favor of new case material.

MODERN CLINICAL PSYCHOLOGY

By T. W. RICHARDS, Northwestern University. *McGraw-Hill Publications in Psychology.* 623 pages, $3.50.

An introduction to the general field of clinical psychology. Integrates material from various fields, such as abnormal psychology and psychiatry, modern testing procedures, current psychodynamic principles regarding the nature of personality and adjustment, with supplementary case material.

FRUSTRATION

The Study of Behavior without a Goal

By NORMAN R. MAIER, University of Michigan. *McGraw-Hill Publications in Psychology.* 264 pages, $3.50.

This important work is a report of an experimental program of research on frustration conducted in the author's laboratory during the past ten years. The volume presents the evidence which leads to a new interpretation of behavior expressed when an individual in a state of frustration. The author departs from the usual approach to abnormal behavior, which seeks to discover how the behavior solves a problem for the individual. Instead, the approach in this book is to indicate that searching for a solution of a problem in the symptom is misleading; rather, according to Professor Maier, symptoms are determined by the principle of availability and are not determined by the end they serve.

Send for copies on approval

McGRAW-HILL BOOK COMPANY, Inc.

330 West 42nd Street — New York 18, N. Y.